

## Volume 77, issue 20, October 2018

65 articles in this issue

# Aims and scope

*Multimedia Tools and Applications* publishes original research articles on multimedia development and system support tools, and case studies of multimedia applications. Experimental and survey articles are appropriate for the journal. The journal is intended for academics, practitioners, scientists and engineers who are involved in multimedia system research, design and applications. All papers are peer reviewed.

Specific areas of interest include (but are not limited to):

## **Multimedia Tools:**

- Multimedia application enabling software
- System software support for multimedia
- Hypermedia
- Performance measurement tools for multimedia
- Multimedia authoring tools
- System hardware support for multimedia
- Multimedia databases and retrieval
- Web tools and applications

## **Multimedia Applications:**

### **+ Prototype multimedia systems and platforms**

- Multimedia on information superhighways

### **+ Home**

- Video on-demand
- Interactive TV
- Home shopping
- Remote home care
- Electronic album
- Personalized electronic journals

### **+ Education and Training**

- Computer aided instruction
- Distance and interactive training
- Multimedia encyclopedias
- Interactive training on the web

# Editors

## Editor-in-Chief:

**Borko Furht**, *Dept. of Computer Science and Engineering, Florida Atlantic University, USA*

## Section Editors:

**Harry Agius**, *Brunel University London, United Kingdom*

**Jungong Han**, *Aberystwyth University, United Kingdom*

**Fabio Narducci**, *University of Salerno, Italy*

**Maria Da Graça Pimentel**, *University of São Paulo, Brazil*

**Weisong Shi**, *University of Delaware, Wilmington, Delaware, USA*

## Associate Editors:

**Sambit Bakshi**, *National Institute of Technology Rourkela, India*

**Carmen Bisogni**, *University of Salerno, Italy*

**Wei Cheng**, *University of Washington, Tacoma, Washington, USA*

**Zheng Dong**, *Wayne State University, Detroit, Michigan, USA*

**Song Fu**, *University of North Texas, Denton, Texas, USA*

**Chiara Galdi**, *EURECOM, Sophia Antipolis, France*

**Dalong Li**, *Torc Robotics, USA*

**Dongfang Liu**, *Rochester Institute of Technology, Rochester, New York, USA*

**Jianping Wang**, *City University of Hong Kong, China*

**Dewei Yi**, *University of Aberdeen, UK*

**Qiang Zhang**, *Xidian University, China*

## Editorial Board:

**Marios C. Angelides**, *Brunel University London, United Kingdom*

**Marco Anisetti**, *University of Milano, Italy*

**Andrew Bagdanov**, *University of Florence, Italy*

**Lamberto Ballan**, *Stanford University, USA*

**Jenny Benois-Pineau**, *University of Bordeaux, France*

**Marco Bertini**, *University of Florence, Italy*

**Bharat Bhargava**, *Purdue University, USA*

**Alberto del Bimbo**, *Università di Firenze, Italy*

**Susanne Boll**, *University of Oldenburg, Germany*

**Ying Cai**, *Iowa State University, Ames, Iowa, USA*

**Aniello Castiglione**, *University of Naples, Italy*

**Pablo Cesar**, *Centrum Wiskunde & Informatica, Netherlands*

**Amit Chakraborty**, *Siemens Corporate Research, USA*

**Shu-Ching Chen**, *Florida International University, Miami, Florida, USA*

**Tat-Seng Chua**, *National University of Singapore*

**Ernesto Damiani**, *University of Milano, Italy*

**Antitza Dantcheva**, *INRIA, France*

**Chabane Djeraba**, *LIFL Laboratory, France*

**Schahram Dustdar**, *Vienna University of Technology, Wien, Austria*

**Abdulmotaleb El Saddik**, *University of Ottawa, Canada*  
**Ben Falchuk**, *Vencore Labs, USA*  
**Marco Furini**, *University of Modena and Reggio Emilia, Italy*  
**Ombretta Gaggi**, *University of Padua, Italy*  
**David C. Gibbon**, *AT&T Labs - Research, USA*  
**Yihong Gong**, *NEC, USA*  
**Sonja Grgic**, *University of Zagreb, Croatia*  
**M. Shamim Hossain**, *King Saud University, Saudi Arabia*  
**Kien A. Hua**, *University of Central Florida, USA*  
**Benoit Huet**, *Institut Eurecom, France*  
**Horace H.S. Ip**, *City University of Hong Kong*  
**Ebroul Izquierdo**, *Queen Mary, University of London, United Kingdom*  
**Ramesh Jain**, *University of California, Irvine, USA*  
**Rongrong Ji**, *Xiamen University, China*  
**Shuqiang Jiang**, *Chinese Academy of Sciences, China*  
**Hari Kalva**, *Florida Atlantic University, Boca Raton, Florida, USA*  
**Shunsuke Kamijo**, *University of Tokyo, Japan*  
**Ahmed Karmouch**, *University Ottawa, Canada*  
**Harald Kosch**, *University of Passau, Germany*  
**Clement Leung**, *Hong Kong Baptist University, Hong Kong*  
**T.D.C. Little**, *Boston University, USA*  
**Huadong Ma**, *Beijing University of Posts & Telecommunications, PR of China*  
**Hong Man**, *Stevens Institute of Technology, Hoboken, New Jersey, USA*  
**Oge Marques**, *Florida Atlantic University, Boca Raton, Florida, USA*  
**Liam M. Mayron**, *Florida Institute of Technology, Florida, USA*  
**Michele Nappi**, *University of Salerno, Italy*  
**Vincent Oria**, *New Jersey Institute of Technology, New Jersey, USA*  
**Charles B. Owen**, *Michigan State University, USA*  
**Dragutin Petkovic**, *San Francisco State University, USA*  
**Giuseppe Polese**, *University of Salerno, Italy*  
**Jorge Posada**, *Vicomtech Research Centre, Spain*  
**B. Prabhakaran**, *University of Texas at Dallas, USA*  
**Martin Reisslein**, *Arizona State University, Tempe, Arizona*  
**Marco Roccetti**, *University of Bologna, Italy*  
**Simone Santini**, *University of California at San Diego, USA*  
**Shin'ichi Satoh**, *National Institute of Informatics, Tokyo, Japan*  
**Klaus Schöffmann**, *Klagenfurt University, Austria*  
**Lorenzo Seidenari**, *University of Florence, Italy*  
**Ali Asghar Nazari Shirehjini**, *Karlsruhe Institute of Technology, Germany*  
**Scott M. Stevens**, *SEI, Carnegie Mellon University, USA*  
**Jinhui Tang**, *Nanjing University of Science and Technology, P.R. China*  
**Yoshiyasu Takefuji**, *Keio University, Japan*  
**Christian Timmerer**, *Klagenfurt University, Austria*  
**Tiberio Uricchio**, *University of Florence, Italy*  
**Jingdong Wang**, *Microsoft Research Asia, China*  
**Meng Wang**, *Hefei University of Technology, China*  
**Qi Wang**, *Northwestern Polytechnical University, China*  
**Lars Wolf**, *Technical University Braunschweig, Germany*

**Changsheng Xu**, *NLPR, Chinese Academy of Sciences, China*

**Hanwang Zhang**, *National University of Singapore, Singapore*

**Zhiyong Zhang**, *Henan University of Science and Technology, China*

**Roger Zimmermann**, *National University of Singapore, Singapore*

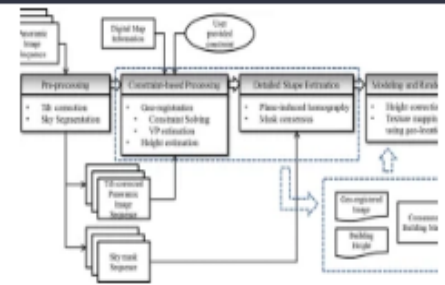
**Ce Zhu**, *Nanyang Technological University, Singapore*

## [Interactive 3D building modeling method using panoramic image sequences and digital map](#)

Hyunki Kim & Soonhung Han

OriginalPaper | Published: 30 March 2018

Pages: 27387 - 27404

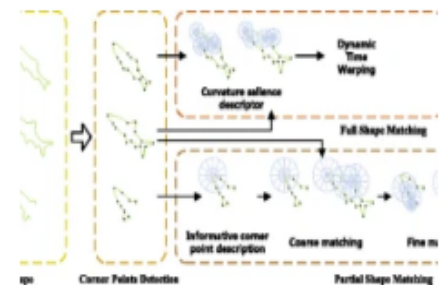


## [A curvature salience descriptor for full and partial shape matching](#)

Zhengbing Wang, Guili Xu ... Zhengsheng Wang

OriginalPaper | Published: 17 April 2018

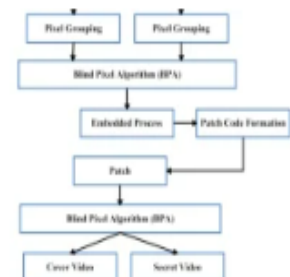
Pages: 27405 - 27426



## [Robust secure video steganography using reversible patch-wise code-based embedding](#)

K. Rajalakshmi & K. Mahesh

OriginalPaper | Published: 31 May 2018 | Pages: 27427 - 27445

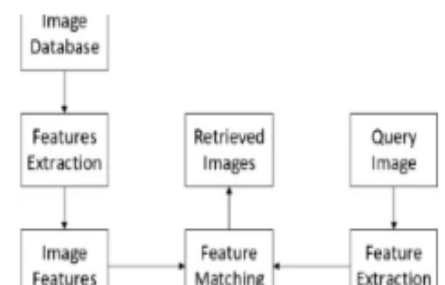


## [Multimedia and geographic data integration for cultural heritage information retrieval](#)

Erasmus Purificato & Antonio M. Rinaldi

OriginalPaper | Published: 03 April 2018

Pages: 27447 - 27469

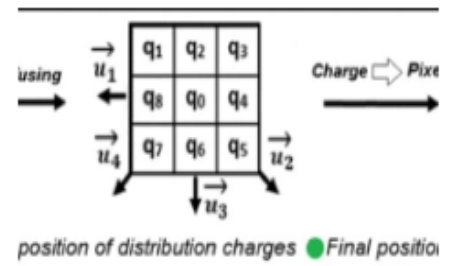


[Face description using electric virtual binary pattern \(EVBP\): application to face recognition](#)

Abdellatif Dahmouni, Karim El Moutaouakil & Khalid Satori

OriginalPaper | Published: 15 April 2018 |

Pages: 27471 - 27489

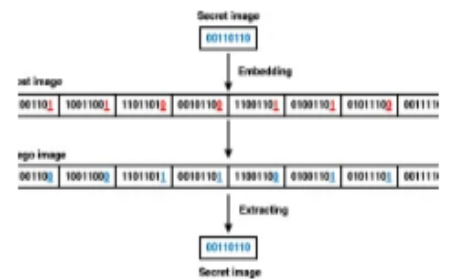


[Artificial bee colony approach for enhancing LSB based image steganography](#)

Anan Banharnsakun

OriginalPaper | Published: 03 April 2018 |

Pages: 27491 - 27504



[The study on increasing the identification accuracy of waxed apples by hyperspectral imaging technology](#)

Huiquan Wang, Haojie Zhu ... Jinhai Wang

OriginalPaper | Published: 04 May 2018 | Pages: 27505 - 27516

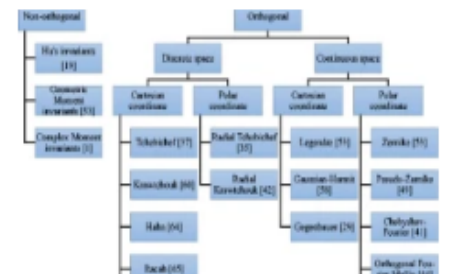


[Efficient 3D object classification by using direct Krawtchouk moment invariants](#)

Rachid Benouini, Imad Batioua ... Hassan Qjidaa

OriginalPaper | Published: 12 April 2018 |

Pages: 27517 - 27542

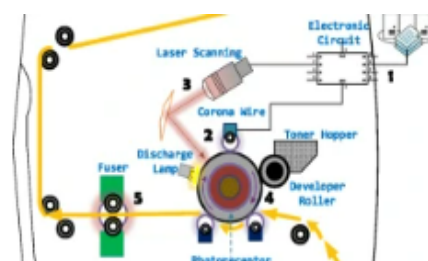


## Decision-theoretic model to identify printed sources

Min-Jen Tsai, Imam Yuadi & Yu-Han Tao

OriginalPaper | Published: 22 April 2018 |

Pages: 27543 - 27587

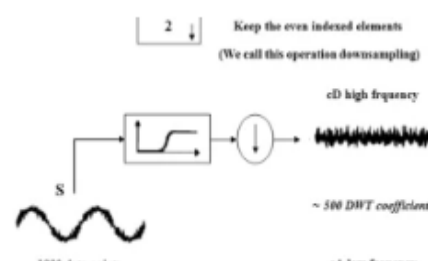


## Blind speech watermarking using hybrid scheme based on DWT/DCT and sub-sampling

Ahmed Merrad & Slami Saadi

OriginalPaper | Published: 21 April 2018 |

Pages: 27589 - 27615



## Low - resolution vehicle recognition based on deep feature fusion

Lixia Xue, Xin Zhong ... Min Hu

OriginalPaper | Published: 04 May 2018 | Pages: 27617 - 27639



## Global motion based video super-resolution reconstruction using discrete wavelet transform

Wasnaa Witwit, Yifan Zhao ... Sri Addepalli

OriginalPaper | [Open Access](#) | Published: 11 April 2018 |

Pages: 27641 - 27660



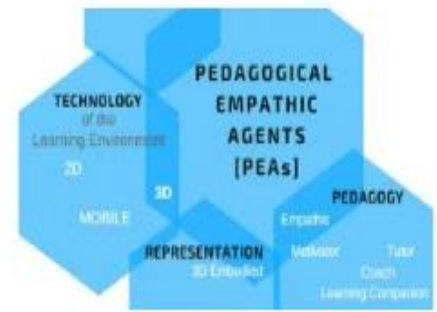


## Architecture and interaction protocol for pedagogical-empathic agents in 3D virtual learning environments

Theodouli Terzidou, Thrasyvoulos Tsiatsos & Hippokratris Apostolidis

OriginalPaper | Published: 18 April 2018

Pages: 27661 - 27684



## Multimodal speaker diarization for meetings using volume-evaluated SRP-PHAT and video analysis

P. Cabañas-Molero, M. Lucena ... N. Ruiz-Reyes

OriginalPaper | Published: 11 April 2018

Pages: 27685 - 27707

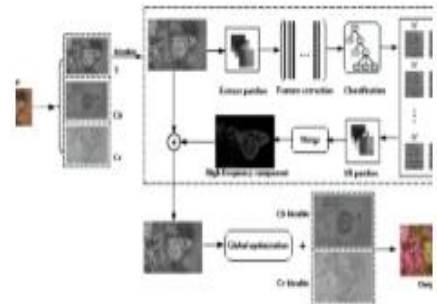


## Super-resolution via supervised classification and independent dictionary training

Ronggui Wang, Qinghui Wang ... Min Hu

OriginalPaper | Published: 14 April 2018

Pages: 27709 - 27732



## Decision-theoretic model to identify printed sources

Min-Jen Tsai<sup>1</sup>  · Imam Yuadi<sup>1,2</sup> · Yu-Han Tao<sup>1</sup>

Received: 9 May 2017 / Revised: 21 February 2018 / Accepted: 26 March 2018 /  
Published online: 22 April 2018

© Springer Science+Business Media, LLC, part of Springer Nature 2018

**Abstract** When trying to identify a printed forged document, examining digital evidence can prove to be a challenge. Over the past several years, digital forensics for printed document source identification has begun to be increasingly important which can be related to the investigation and prosecution of many types of crimes. Unlike invasive forensic approach which requires a fraction of the printed document as the specimen for verification, noninvasive forensic technique uses the optical mechanism to explore the relationship between the scanned images and the source printer. To explore the relationship between source printers and images obtained by the scanner, the proposed decision-theoretical approach utilizes image processing techniques and data exploration methods to calculate many important statistical features, including: Local Binary Pattern (LBP), Gray Level Co-occurrence Matrix (GLCM), Discrete Wavelet Transform (DWT), Spatial filters, the Wiener filter, the Gabor filter, Haralick, and SFTA features. Consequently, the proposed aggregation method intensively applies the extracted features and decision-fusion model of feature selections for classification. In addition, the impact of different paper texture or paper color for printed sources identification is also investigated. In the meantime, the up-to-date techniques based on deep learning system is developed by Convolutional Neural Networks (CNNs) which can learn the features automatically to solve the complex image classification problem. Both systems have been compared and the experimental results indicate that the proposed system achieve the overall best accuracy prediction for image and text input and is superior to the existing approaches. In brief, the proposed decision-theoretical model can be very efficiently implemented for real world digital forensic applications.

**Keywords** Decision fusion · Scanner · Feature filters · Feature selection · Support Vector Machines (SVM) · Deep learning · Convolutional Neural Networks (CNNs)

---

✉ Min-Jen Tsai  
mjtsai@cc.nctu.edu.tw

<sup>1</sup> Institute of Information Management, National Chiao Tung University, 1001 Ta-Hsueh Road, Hsin-Chu 300, Taiwan, Republic of China

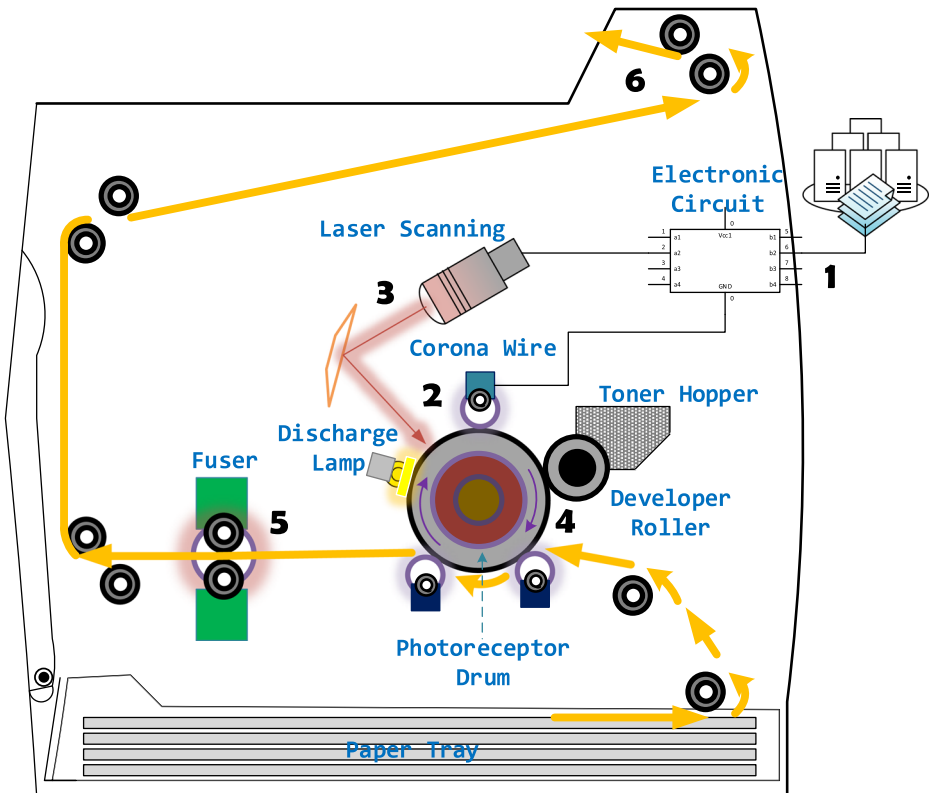
<sup>2</sup> Department of Information and Library Science, Airlangga University, Jl. Airlangga 4-6, Surabaya, East Java 60286, Indonesia

## 1 Introduction

Digital forensics is the examination and analysis of digital evidence to prove the occurrence of a crime. Digital forensics implements specific tools and methods to identify, collect, and analyze digital evidence [27]. Recently, digital forensics for printed document source identification has begun to be increasingly important in the investigation and prosecution of many types of crimes. The challenges in the field of forensic investigation still rise to provide appropriate and sufficient security measures and devices [1, 22] in the forensic process to help forensic investigation. However, these devices give the potential effects where the digital documents often contain information about crimes committed, movement of suspects, and hidden messages. Correspondingly, documents in a suspect's possession might possibly reveal clues from digital evidence. Using digital evidence in a legal trial can prove to be challenging. Therefore, it requires accurate techniques to prove authenticity of digital evidence. Digital forensics experts are needed to assist law enforcement to determine whether a suspect is guilty or innocent of a crime, by examining evidence using standard investigation techniques as well as a broad range of digitizing tools such as cameras, scanners, and microscopes.

Essentially, each printer produces printed document that has distinctive texture characteristics compared to the others. As printers are electromechanical devices with moving parts, there are many small physical differences on printers such as motor drifting and gear precision that can be seen on printed pages. These information patterns can be used as intrinsic signatures of these devices [36]. The studies [33–35] discussed the major banding artifact from the printed output due to quasi-periodic fluctuations of the printer. These are primarily due to fluctuations in the angular velocity of the photoreceptor drum (shown in Fig. 1) and result in non-uniform scan line spacing. The main cause of banding is electromechanical fluctuations in the printer mechanism, mostly from gear backlash. Because these fluctuations are related to the gearing, the banding frequencies present in the printed page should directly reflect the mechanical properties of the printer. It is described as non-uniform light and dark lines perpendicular to the direction in which the paper moves through the printer. This causes a corresponding fluctuation in developed toner on the printed page.

Printed documents generated from different laser printers will produce differences of texture structure either for printed text or images on the paper, in general with a toner cartridge. Toner particles are specifically melted by the heat of the fuser, and are thus bonded to the paper. The intrinsic signatures from the printer, which can be detected on a document paper including shapes, sizes and patterns can be used as a guide for researchers to distinguish and classify the printer sources [1, 8, 33–36, 53, 55, 56]. There are several approaches for authenticating printed documents and Table 1 illustrates the literature papers on the associated topics for source identification. Mikkilineni et al. [33–36] applied GLCM for each English “e” character to form the feature vectors. In their experiment [33], two strategies were developed for printer identification based on examining the printed document. Finding intrinsic signatures of the printed document was the first solution for identifying the characteristics of a particular printer, model and manufacturer's brand with very high resolution (2400 dpi). Meanwhile, the other one was banding by detecting the extrinsic signature with embedded information from a document with electrophotography (EP) printers in modulating the intrinsic feature. They implemented a 5-Nearest-Neighbor (5NN) classifier in their previous work [33, 34] and in other experiments [35, 36], SVM was applied to classify 10 printer types. Several researchers in [53, 55, 56] also conducted their experiments by using GLCM with different approaches with extended features. Tsai et al. [53, 55] implemented GLCM and DWT based feature extraction to identify Chinese character and used feature selection to get the optimum feature



**Fig. 1** Basic laser printer components and operation

set for printer source identification. In further study [56], they identified Japanese character with more features which include GLCM, DWT, Gaussian, LoG, Usharp, Wiener and Gabor features for classification. Furthermore, Kee and Farid [21] proposed two solutions using principal component analysis (PCA) and singular value decomposition (SVD) for printed characters to distinguish source printers. Wu et al. [58] extracted the geometric distortion of Chinese document as the intrinsic features for classification.

On the other hand, examinations of intrinsic marks for image document were conducted with various techniques. Choi et al. [5] used the noise features extracted from the statistical analysis of the HH sub-band of DWT for 15 RGB channel features and 24 Cyan, Magenta, Yellow and Black (CMYK) channel features for identifying the source of color laser printer. Alternatively, Kim and Lee [23] applied each CMY color channel in the discrete Fourier transform (DFT) domain to identify the color laser printer. Accordingly, Ryu et al. [45] investigated the property of halftone textures from electro photographic printer in each channel of CMYK domain. They applied the Hough transform and constructed the histogram by angle values. Correspondingly, Bulan et al. [3] also assessed the similarity of a pair of geometric distortion signatures during the printing process using the normalized correlation.

According to the above mentioned methods, the use of feature filters and SVM classification are the most commonly adopted solutions. Nevertheless, the use of feature sets are generally implemented independently based on the limit of computation complexity. As a consequence, the feature filter applicability of the underlying techniques involved can still be improved. For example, how

**Table 1** Research papers on the topic of identifying printed document

Document type	Research	Approach	Research object	Classifier	Claimed average accuracy rate	Number of printer
Text document	[35]	GLCM	English Character “e”	SVM	93.0%	10
	[55]	GLCM, DWT	Chinese character “永”	SVM	98.64%	12
	[56]	GLCM, DWT, Gaussian, others	Japanese character “シ”	SVM	94.23%	12
	[8]	GLCM_MDMS, GLCM_MD, CTGF_GLCM_MDMS and others	English Character “e” and frame document	SVM	98.47%	10
Image Document	[5]	DWT, GLCM	Color image document	SVM	88.75%	8
	[23]	Discrete Fourier transform (DFT)	Photograph Image document	SVM	94.4%	7
	[45]	Hough transform	Photograph Image document	a	91.9%	9

<sup>a</sup> denotes no exact information provided

to wisely select the most important features among large feature sets is still a critical issue. Thus, to accomplish better accurate identification rate is technically inspiring for researchers to enhance the existing techniques and explore new forensics approaches.

To achieve those goals, the objective of this paper is to obtain the best performance of printer source identification for text and image documents where we investigate different filter sets, expand feature filter combination, and theoretically diffusing the feature selection among feature space. Furthermore, the aim of the research is to acquire the best decision results.

In addition, the up-to-date deep learning technology of artificial intelligence has been adopted lately in many different fields. The advancement of the approaches not only reduces human intervention but also learns the model by the input data automatically. In the meantime, this study will systematically compare the performance for feature based SVM system and deep learning based classification system for text and image documents in details. As a consequence, this paper is structured as following: Section 2 reviews the theoretical background of different statistical approach for feature representation, feature selection, proposed identification scheme and deep learning based classification system. Section 3 illustrates the experimental results with comparison and discussion. Section 4 concludes the paper and discusses the future works.

## 2 The theoretical background of feature statistics and research approach

Before the introduction of investigated features and research approach, we briefly review the functional mechanism of the laser printers.

### 2.1 How laser printer works

Laser printer is an electro photographic printer that uses a focused beam or light to transfer text and images onto paper. The printer process varies among different manufacturers and the

printed document is greatly influenced by the printer mechanism [16]. Figure 1 illustrates the steps of basic laser printing process.

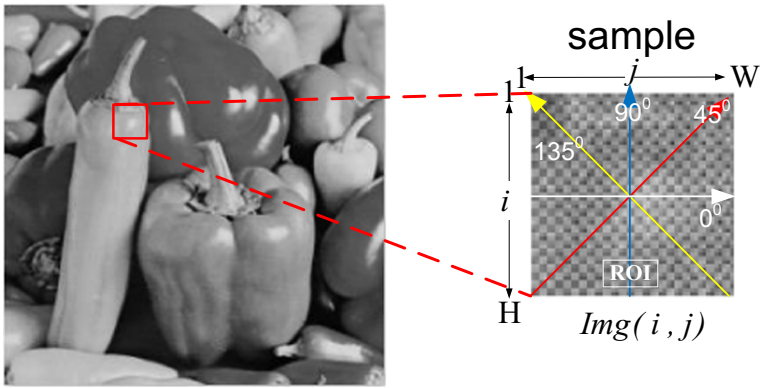
- 1) The data flow of text or image documents are fed into the printer from a computer.
- 2) An electronic circuit with a high-voltage wire gives a static electric charge to activate the corona wire. It gives a total positive charge to the photoreceptor drum.
- 3) At the same time, a tiny laser shines across the drum surface to discharge certain points. It draws a pattern of electrical charge (an electrostatic image) printed as character or image pattern. This system works with a positive electrostatic image on a negative background.
- 4) After the pattern is set, a developer roller fetches powder particles from tonner hopper and attaches to the negative discharged areas the photoreceptor drum to coat it onto a paper with positively charged toner.
- 5) The paper which has already revolved along with the drum and toner hopper passes through two hot fusers.
- 6) Finally, the printout is completed and come out from the other side of the printer [17]

## 2.2 Feature filters

Features are important characteristics extracted from the printed documents and will be analyzed for classification purpose. In this study, ten different set of filters are adopted to acquire the most informative values for analysis. Due to the limit of the space, concise description is explained for these filters and the formulas are tabulated in the [Appendix](#) section. Interested readers can refer [2, 4, 6, 7, 9–12, 28, 30, 32, 38, 39, 47, 54, 57, 59] for more detailed information.

### 2.2.1 The GLCM features

GLCM is a popular statistical method for texture recognition that provides the spatial relationship of the pixels. Four directions are used to generate the data that are focused on the generation of the matrix, i.e., 0 degree (horizontal direction), 45 degree direction, 90 degree (vertical direction), and 135 degree direction. The direction and spatial distance from the reference pixel  $i$  will be defined, such as 1 space at 45 degrees direction locates at the adjacent pixel  $j$  next to the reference pixel  $i$  [34, 54, 55]. The features are the estimation of the second order probability density function of the pixels in the image and GLCM is a matrix of frequencies, where each element  $(i, j)$  indicates the spatial location of image. The  $g_{lcm}(n, m)$  means the number of occurrences of pixels with gray levels  $n$  and  $m$  respectively with a separation of  $(dr, dc)$  pixels and the number of rows and columns is determined by the number of grayscale intensity values in the image. If the GLCM is normalized with respect to  $R_{g_{lcm}}$ , its entries then represent the probability of occurrence of pixel pairs with gray levels  $n$  and  $m$  with separation  $(dr, dc)$ . Here we choose  $dc=0$  and  $dr=1$ . A binary image map with all the pixels labeled as 1 within region of interest (ROI), while pixels valued as 0 if they are not within ROI. The number of pixels is in the ROI as shown in Fig. 2, which is the set of all pixels within the printed area of the image. With this intention, the formulas of ROI and the estimated values of normalized GLCM are listed in the [Appendix](#). There are a total number of 22 different textural features that could be computed from the GLCM, such as described in [34].



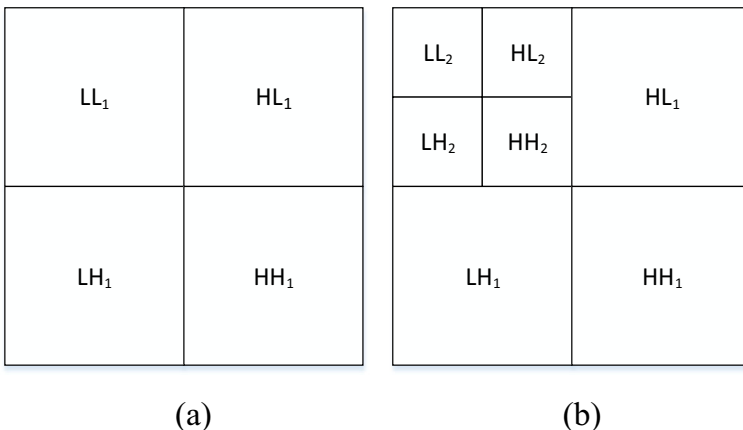
**Fig. 2** An example of ROI for printed pepper and the four different orientations for generation of GLCM

2.2.2 The wavelet-based feature

The DWT feature set in this study focused on a two-dimensional scaling wavelet that is a product of two one-dimensional functions based on the research of [9, 54, 55]. There are four subbands which are sub-sampled as decomposed image shown in Fig. 3. The sub-bands can give a label LH<sub>1</sub>, HL<sub>1</sub> and HH<sub>1</sub> and the detailed image LL<sub>1</sub> corresponds to the coarse level coefficients such as an approximation image. However, the sub-band LL<sub>1</sub> is critical sampled after decomposing. This process is a two-level wavelet decomposition and can be further decomposed by using LL<sub>2</sub> until the final scaled decomposition is accomplished. By using this filter set, there are 12 statistical features which include standard deviation, skewness, and kurtosis for four HH, LH, HL, LL sub-bands [9, 10, 28, 54].

2.2.3 The spatial features

The spatial filters applied in this study are Gaussian smoothing filter, Laplacian of Gaussian filter, Unsharp, Gabor, and Wiener filters. Those filters are briefly explained as following:



**Fig. 3** DWT Image decomposition (a) One-level (b) two levels

- 1) By using the process of convolution, Gaussian filter can minimize the image noise based on probability distribution function [4]. In this study, Gaussian smoothing is implemented to enhance image (text and image document) structures at different scales.
- 2) The second derivative of Gaussian filter is Laplacian of Gaussian (LoG). It is a filter function derivative which can detect areas with rapid change as edge on the image, combining noise reduction and responding to changes in image gradients [32]. Laplacian is very vulnerable or sensitive to the presence of noise.
- 3) The unsharp filter is a high pass filter with a sharpening operator which enhances edges through the procedure which subtracts an unsharp, or smoothed version of an image from the original image. Implementing unsharp filter is to reduce blur in scaled unsharp version of the image from the original image. In practice, it can be done by subtracting the blurred image from the original image [47, 59]. This study used unsharp filters with the high-frequency Laplacian filter to retrieve the numeric information from the scanned text and image document.
- 4) The 2-D Gabor filters are sinusoids modulated by a Gaussian window that has several advantages such as invariance to illumination, rotation, scale, translation [7, 57]. In the spatial and the spatial-frequency domain, it is a set of orientation and frequency sensitive band pass filters which have the optimum joint resolution. These filters are suitable for extracting orientation dependent frequency contents of patterns and have been widely used in solution for forensics identification. This study implemented the Gabor orientation for scanned image upon four values, i.e.  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  and  $135^\circ$ . The variance along  $x$  and  $y$  axis,  $f$  is the frequency of sinusoidal function and  $\theta$  is the orientation of Gabor function. The settings of  $s_x = 4$ ,  $s_y = 12$ ,  $f = 4, 12, 24, 48$ ,  $\theta = 0, \pi/4, \pi/2$ , and  $3\pi/4$  yield 48 different feature filters for feature extraction [11].
- 5) The Wiener filter was introduced by Norbert Wiener for minimizing the noise impact of the image by additive noise and blurring. It is a technique based on statistical approximation from the local neighborhood of every pixel and forms the foundation of data-dependent linear least square error filters. Additionally, it plays a central role in a wide range of applications such as channel equalization, linear prediction, signal restoration, echo cancellation, system identification and channel equalization. The filter can be applied in frequency domain as an adaptive linear filter that works on local variant characteristics of an image. It also has good performance on removing the Gaussian white noise [57]. By using these filters, there are 64 features in this category.

#### 2.2.4 The Haralick features

The other spatial feature of co-occurrence matrix and texture features are Haralick texture features. It was formulated by R. M. Haralick in 1973 by computing various statistical properties that was used to construct the matrix obtained by using the directions of  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  and  $135^\circ$  [12]. However, the spatial grayscale level reliance matrix at the direction and spatial distance such as GLCM feature filter where  $(i, j)$  indicates the spatial location of image;  $g_{lcm}(n, m)$  means the number of occurrences of pixels with gray levels  $n$  and  $m$  respectively with a separation of  $(dr, dc)$  pixels. Each element represents as a sum of total number of pairs from gray levels at the position in the predefined offset over the full image. Accordingly, it has fourteen features of co-



occurrence matrix on the spatial gray level dependence. The fourteen functions are calculated from each image in the matrices as feature filters before implemented in classification and the details of the 14 features are defined in the [Appendix](#).

### 2.2.5 The fractal features

In this study, we extracted fractal based features by calculating the fractal dimension. These features are built on fractal dimension for gray-scale images which depict objects and structure boundary. Three primary solutions are applied in texture analysis for fractal feature values. In the first solution, an image surface is generated by using model of a fractional motion and its roughness as a parameter for discrimination between classes. In the second one, a modeled image is generated as an intensity surface in its fractal dimension as the texture parameter. And the last solution, making an image as a point set union forms a vector for totally fractal dimension. More importantly, the fractal filters not only computes the fractal dimension of any grayscale image but measures the roughness of the image with different textures. The extraction algorithm consists in decomposing the input image into a set of binary images from which the fractal dimensions of the resulting regions are computed in order to describe segmented texture patterns. The fractal feature vector [6] -  $\Delta(x, y)$  is defined in the [Appendix](#). Costa [6] and Bekhti [2] implemented threshold set of binary ( $n_t$ ) corresponding to the maximum possible gray level in  $I(x, y)$  images which yield 48 features. To extract features, they used  $F = (I, nt)$  where  $nt$  represents a threshold and we used 4 as the threshold that acquired 24 features.

### 2.2.6 The LBP features

LBP is a feature extractor that has an appropriate and powerful sub pattern-based texture descriptor. It characterizes the gray-scale invariant texture and combination between measuring texture from each neighborhood and the difference of the average gray level of those pixels based on binary numbers. It has the advantage to obtain pattern labels of image texture by calculating histogram of uniform LBP, especially for texture with different edges and shapes [30, 38, 39]. LBP operator defined in the [Appendix](#) efficiently supervises texture segmentation and is used together with the size of the local contrast as a feature filter with high quality performance.

In this study, the sample image is divided into blocks for the LBP feature extraction. LBP (8,1) neighborhood is used with the pixel in the printed area which compared the pixel to each of its 8 neighbors that the direction is from left-top, left-middle, left-bottom, right-top, etc. It is done for each pixel in a block pixel when the pixel value is greater than the neighbor's value, write "0". Otherwise, write "1" as the LBP value. Finally, the histogram value from each block is computed as combinations of which pixels are smaller or greater than the center pixel value. The total value of LBP are merged into one histogram when using uniform patterns and the length of feature vector reduced from 256 to 59 for a single cell if  $R=1$  and  $P=8$ . The combined 59 features are further used as feature filters in this study.

## 2.3 Support vector machine

SVM is a classifier that can generalize texture patterns in a high-dimensional space and classify pixels for images according to textural cues. It is a concept for classification, regression, and other learning tasks that can be simply explained as an attempt to find the best hyperplane which serves as a separator among classes in the input space. The best hyperplane separation among them can be found by measuring the margin hyperplane and looking for maximum points [24]. In addition, SVM is able to obtain the best result in comparison among feature extractions for a multi texture classification problem [52]. The SVM generates a model based on the training data and can predict the target values of the test data given only the test data attributes [15, 52]. Given a training set of instance-label pairs  $x_i, y_i, i = 1, \dots, l$  where  $x_i \in R^n$  and  $y \in \{1, -1\}$ , the SVM classification can be formulated in Eq. (1).

$$\min_{w,b,\xi} \frac{1}{2} w^T w + C \sum_{i=1}^l \xi_i \quad \text{subject to } y_i(w^T \phi(x_i) + b) \geq \xi_i, \xi_i \geq 0 \quad (1)$$

where  $C$  is the capacity constant,  $w$  is the vector of coefficients,  $b$  is a constant, and  $\xi_i$  represents parameters for handling non separable data (inputs). The kernel  $\phi$  is used to transform data from the input (independent) to the feature space. The functional dependence of the dependent variable  $y$  on a set of independent variable  $x$ . Note that  $y \in \pm 1$  represents the class labels and  $x_i$  represents the independent variables. However, the index  $i$  labels the  $N$  training cases. Here training vectors  $x_i$  is mapped into a higher (maybe infinite) dimensional space by the function  $\phi$ . SVM finds a linear separating hyperplane with the maximal margin in this higher dimensional space.  $C > 0$  is the penalty parameter of the error term. Furthermore,  $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$  is called the kernel function. The radial basis function (RBF)-based kernel function builds the classifier for this study [15] as defined in Eq. (2) where  $\gamma$  is a parameter that sets the “spread” of the kernel.

$$K_{RBF}(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right), \gamma > 0 \quad (2)$$

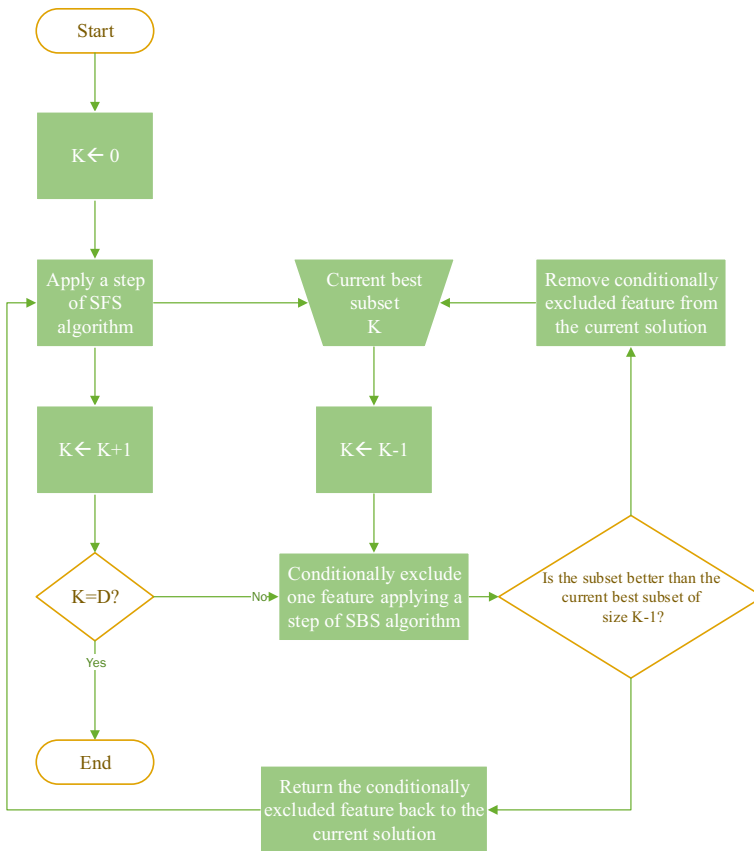
The images that have been extracted in different features are then classified by using SVM utilities to get the optimized parameters. The Optimal kernel parameter for  $C$  and  $\gamma$  were obtained by a coarse grid search in the parameter space within the interval  $C$  and a mapping  $\phi$  is considered to transform the original data space into another feature space.

## 2.4 Decision-fusion model

As described in Sec. 2.2, as more features are adopted, the complexity of computation and processing time will also increase for classification. Therefore, it is an important issue to identify the critical features among the feature sets in order to reduce the overall computing time and complexity. The decision-theoretical model referring to the feature-selection aggregation for decision technique is therefore explored. Several feature

selection methods will be utilized and the floating search methods are implemented by the sequential selection procedures that are related to the *plus l take-away r* algorithms [40]. It consists of applying Sequential Forward Selection (SFS) for  $l$  times and then followed by  $r$  steps of Sequential Backward Selection (SBS) with its fixed cycle of forward and backward selection repeated until the required number of features is reached.

SFS is a greedy search algorithm when it starts from the empty or blank empty set. The next step is sequentially add the feature  $x^+$  which maximizes  $J(X_k + X)$  when combined with the features  $Y_k$  that have already been selected. On the other hand, SBS works in the opposite direction of SFS when it starts from the full set ( $Y_0 = X$ ) and sequentially remove the worst feature  $x^-$  that can reduces the value of the objective function  $J(Y - X^-)$ . By removing a feature, it possibly increases the objective function  $J(X_k - X) > J(Y_k)$ . The sequential forward and backward selection is updated whenever the modification results achieve better performance. Based on the idea of the SFS ((1,0)-



**Fig. 4** Flowchart of applying SFS and followed by SBS method

search) and SBS ((0,1)-search) algorithms, it can be efficiently implemented by considering conditional inclusion and exclusion of features controlled by the value of the criterion itself. A flowchart of the SFS algorithm is shown in Fig. 4. In this flowchart, the SFS and the SBS algorithms are implemented [29, 42]. Plus  $l$  minus  $r$  selection (LRS) starts from the empty set and repeatedly adds  $l$  features and removes  $r$  features when  $l$  is more than  $r$ . Conversely, when  $l$  is less than  $r$ , LRS starts from the full set and repeatedly removes  $r$  features followed by  $l$  additions.. The *plus  $l$  take-away  $r$*  algorithms method can be described in an algorithmic way as following:

*Input:*  $Y = \{y_j \mid j = 1, \dots, D\}$  //available measurements//

*Output:*  $X_k = \{x_j \mid j = 1, \dots, k, x_j \in Y\}, k = 0, 1, \dots, D$

*Initialization:* if  $l > r$  then  $k := 0$ ;  $X_0 := \emptyset$ ; go to Step 1  
 else  $k := D$ ;  $X_D := Y$ ; go to Step 2

Step 1 (*Inclusion*)

repeat  $m$  times

$$x^+ := \arg \max_{x \in Y - X_k} J(X_k + x)$$

$$X_{k+1} := X_k + x^+; k := k + 1$$

go to Step 2

Step 2 (*Exclusion*)

repeat  $r$  times

$$x^- := \arg \max_{x \in X_k} J(X_k - x)$$

$$X_{k-1} := X_k - x^-; k := k - 1$$

go to Step 1

It can be implemented by using plus 2 minus 1 (P2M1) where ( $l=2, r=1$ ), plus 3 minus 2 (P3M2) where ( $l=3, r=2$ ), and plus 4 minus 3 (P4M3) where ( $l=4, r=3$ ). Furthermore, to perform feature selection, Pudil et al. [41] proposed the SFFS and SBFS methods. The SFFS method is a modified plus- $m$ -minus- $r$  by one more mechanism in the minus step. To ensure for finding the best feature subset, it can repeat the removal step and SBFS method has the same principle as SFFS sequence method. The Sequential Floating Forward Selection (SFFS) starts from the empty set. It works after each forward step and then performs backward steps as long as the objective function increases. Otherwise, the Sequential Backward Floating Selection (SBFS) starts from the full set. After each backward step, it performs forward steps as long as the objective function increases. The SFFS method can be described algorithmically in a

similar way to the previous method as following:

*Input*  $: Y = \{y_j \mid j = 1, \dots, D\}$  //available measurements//

*Output*:  $X_k = \{x_j \mid j = 1, \dots, k, x_j \in Y\}, k = 0, 1, \dots, D$

*Initialization*:  $X_0 := \emptyset; K := 0$

(in practice one can begin with  $k = 2$  by applying SFS twice)

*Termination*: Stop when  $k$  equals the number of features required

Step 1 (*Inclusion*)

$$x^+ := \arg \max_{x \in Y - X_k} J(X_k + x)$$

$$X_{k+1} := X_k + x^+; k := k + 1$$

Step 2 (*Conditional Exclusion*)

$$x^- := \arg \max_{x \in X_k} J(X_k - x)$$

if  $J(X_k - \{x^-\}) > J(X_{k-1})$  then

$$X_{k-1} := X_k - x^-; k := k - 1$$

go to Step 2

else

go to Step 1

A challenge of feature selection integration or fusion represents the method of combining the above mentioned five different techniques of feature selection (P2M1, P3M2, P4M3, SFFS, and SBFS). The goal here is to gather the most useful features from all the selection methods, in such a way that the end-result is to achieve the best outcomes from each technique respectively and then making a fusion from each of them after aggregation.

This study considers decision-theoretical model as experts and feature sets as alternatives. It gives each feature the ranking order or confidence level in the final subset since features in an optimal subset by selection algorithms are generated in a series of inclusion and exclusion steps. Consequently, the count-based aggregation is selected as the algorithm of decision fusion. The feature will get a recommending label whenever a feature is chosen into the optimal subset by a selecting algorithm. Based on the majority vote, the features with the most labels are selected into the final optimal subset. The completed selection is taken by the majority of the features of each group. The highest-level decision or the top  $\lambda$  fusion operates directly on classification applied on the basis of each selection. Thus, the final decision is taken by combining the majority vote from the feature sets.

## 2.5 The proposed approach

We propose the technique based on machine learning method by examining the text and image documents. The diagram shown in Fig. 5 illustrates the identifying procedures which can be

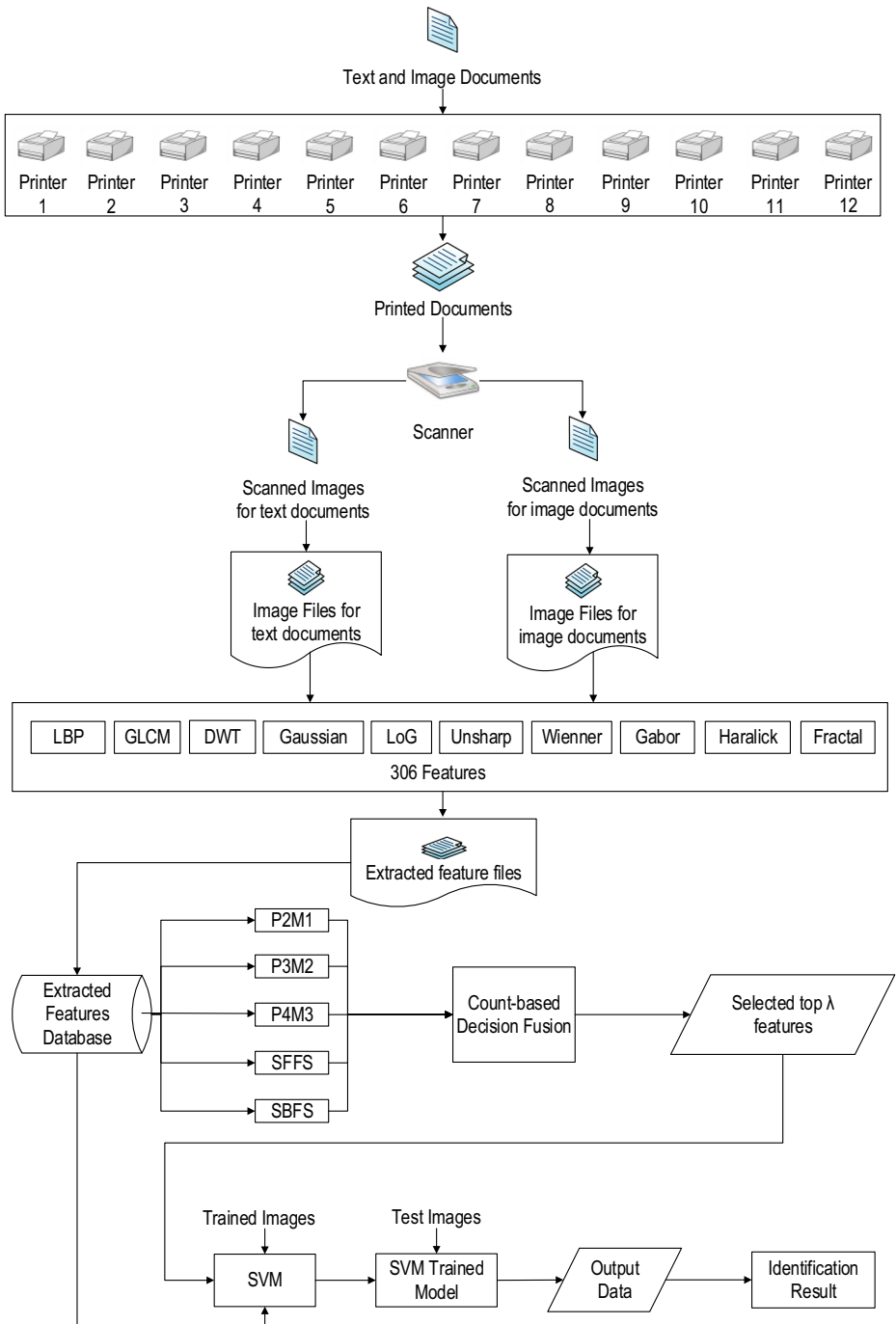


Fig. 5 The research flowchart of SVM based system

divided into five stages such as, printing documents, digitizing documents, feature extraction, feature selection, and classification:

- (1) *Printing documents*: First of all, text and image documents are prepared. For example, different font-type with different font size for characters are studied. In this paper, several printed characters are identified such as commonly used English character “e”, Chinese character “永”, Arabic character “ﻉ”, and Japanese character “シ”. We also investigated the image documents (i.e. Lena, Peppers, Baboon and Wikipedia images) for comparison. Next, all the documents are printed by using 12 printers where brands and models are shown in Table 2.
- (2) *Digitizing documents*: After all the documents have been printed, the second step is to digitize document by using HP Scanjet G4050 with 300 dpi. At this stage, document types are recognized as either text or image document. If the document is identified as text file, region of interest for the extracted characters will be further labelled to remove the blank space in order to obtain the most valuable scanned information. The digital documents are generated in bmp format and cut automatically by Netbean IDE 8.0 software to facilitate the process. For example, each 10 pt. text image is then cropped into  $51 \times 51$  pixel dimension for each identified character that represents the printer source document.
- (3) *Feature extraction*: Extracting the grayscale documents by the proposed filters such as LBP, GLCM, DWT, Gaussian, LoG, Unsharp, Wiener, Gabor, Haralick, and fractal filters. Each character or images are extracted by using ten different feature sets based on the printer sources. For example, the character “e” that originates from each type of printer, we extracted at least 1200 images for each printer by using different feature filters into numeric values. In this study, there are total 306 statistical features which is a very large feature space compared with features adopted in [1, 3, 5, 8, 21, 23, 33–36, 45, 53, 55, 56, 58]. Therefore, the computational complexity is also a critical issue to be resolved in real applications. Therefore, feature selection will be conducted in the next stage to alleviate the computation demands.
- (4) *Feature selection and fusion*: The adaptive feature selection algorithm is implemented here in order to find the most important  $\lambda$  features which helps to reduce the total evaluation time without the loss of accuracy. Five feature selection algorithms P2M1, P3M2, P4M3, SFFS, and SBFS are adopted for the feature selection processes [55]. The count-based aggregation is utilized at decision fusion stage where each feature will get a recommending label whenever a feature is chosen into the optimal subset by a selecting

**Table 2** Laser printer brand and models used in this study

No	Brand	Model
s1	Avision	AM/MF 3000
2	HP	LaserJet Pro 200 Color P. M251nw
3	HP	LaserJet Pro 500 MFP M570dn
4	HP	Color LaserJet CP3525
5	HP	LaserJet Pro CP1025
6	HP	LaserJet 4300
7	HP	LaserJet 4200dtn
8	HP	LaserJet M1132 MFP
9	HP	LaserJet Pro 400 MFP M425dn
10	HP	LaserJet M1522nf
11	HP	LaserJet Pro M1536dnf
12	OKI	C5950

algorithm. Based on the majority vote, the features with the most labels are selected into the final optimal subset which will be used for classification.

- (5) *Printer classification*: The last step of the source identification is to classify the printed sources from different printers using the optimal features from step (4) by using SVM trained model. The extracted images that have been in the numeric value are then inserted into MySQL database. The database contains different schema and query based on printed document type which will be evaluated. Afterwards, this study classifies them by using SVM in the Java environment (Eclipse Indigo) and same SVM parameters applied in [33, 53, 56] are adopted here where 500 images for training and 300 images for testing.

## 2.6 Machine learning and deep learning

In the past, based on professional knowledge and experience, features are extracted heavily depending on human involvement. The process of manually feature selection is usually laborious and time consuming. Therefore, machine learning is a discipline that specializes in how computer simulates and realizes human learning behavior. If the technique can program computer to automatically learn the characteristics and speed up the whole procedures, it can save dramatic amount of time and money.

Accordingly, it is important to understand how the human brain works, and Hubel et al. [18] had found that the operation of the neuron system is hierarchical based on the functional analysis of the cortex cells of the cat to find the corresponding relationship between neurons. Under such understanding, the integration of artificial neural network and the back propagation facilitate the analysis of a large number of input training samples to get the statistical regularity, and to make the prediction.

Hinton et al. [14] have reported that the multi-hidden artificial neural network has excellent learning capability. The learned features are conducive to visualization and classification. In addition, deep neural networks can overcome the difficulty of training through “layer-wise pre-training”. Consequently, deep learning can establish a multi-hidden layer of learning model to process a large number of training samples, learn useful features and enhance the classification or prediction accuracy. Deep learning has also developed many different models such as: Auto Encoder, Sparse Coding, Restricted Boltzmann Machine (RBM) [31], Deep Belief Networks and Convolutional Neural Networks (CNN). Related studies have achieved great success in many fields of human machine interaction [25, 26]. Since many advanced and complicated deep learning structures are generally based on CNN, this study will implement CNN of [25] for comparison purpose.

Convolutional neural networks (CNNs) is a type of artificial neural network that was proposed by Lecun et al. [26], it has become one of the popular tools in the field of speech analysis and image recognition. Its shared weights network structure is very similar to the actual biological neural network in simulation. This feature can also reduce the complexity of network model and reduce the number of parameters. CNNs can directly use the image as the input which can avoid the traditional identification method by complex feature extraction and data reconstruction. The network structure of a convolutional neural network is highly invariant for image translation, scaling up/down, tilting or other forms of deformation.



Convolution neural network [25] is a multi-layer neural network that each layer consists of multiple two-dimensional planes. This plane is composed of multiple independent neurons, the basic concept of network architecture as shown in Fig. 6. The C layer in the graph is the feature extraction layer, and the input of each neuron connects with the local receptive field of the previous layer, and the local feature is extracted. When the local feature is extracted, its positional relationship with other features will also be determined. The S layer in the graph is the feature mapping layer, and each feature layer of the network is composed of multiple feature mapping layers. Each feature is mapped to a plane, and the weights of all the neurons on the plane are equal. The feature mapping structure adopts the influence function kernel, the small sigmoid function acts as the activation function of the convolution network, so that the feature map has the displacement invariance. In addition, since neurons on a map surface share weights, it is possible to reduce the number of network parameters and reduce the complexity of network parameter selection. Each feature extraction layer (C layer) in the convolutional neural network is followed by a computational layer (S layer) for local averaging and secondary extraction.

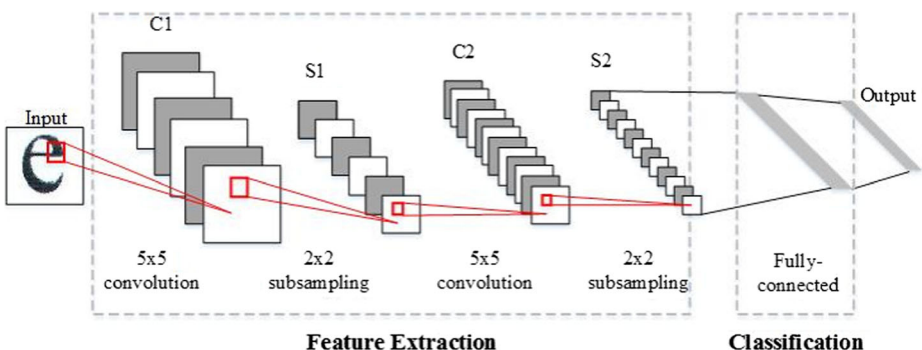
Convolution neural network consists of five basic network layers: input layer, convolutional layer, rectified linear unit (ReLU), pooling layers, and fully connected layer:

(1) Input Layer:

The entrance for the input data.

(2) Convolutional Layers:

This layer contains a series of fixed-size filters, which are used to operate the convolution of the input data, resulting in the so-called eigenvalue map (feature map). These filters can provide useful modules for image recognition, such as image edges, regular patterns, and color changes. The amount to be used depends on the size of the data, the complexity of the image, and the size of the image. There are two important settings, Padding is set to determine how many columns around the image to increase or how many rows of null pixels added; and Stride is set in the filter about how many pixels to move during the scan. These two parameters determine how the filter can scan the full image



**Fig. 6** The network architecture of CNNs

### (3) Rectified Linear Unit (ReLU):

ReLU generally follows the operation of the convolutional layer and provides the output with the non-saturating activation function  $f(x) = \max(0, x)$ . According to Krizhevsky's study [25], these equations can be used in convolution of neural networks in the rapid convergence of training, also dealing with the elimination of gradient problems to speed up the training.

### (4) Pooling Layers:

The pooling layer is to reduce the spatial size of the representation and the number of parameters. It also minimizes the computation number through the feature map and controls overfitting. The pooling operation arranges several forms of translation invariance and sporadically inserts the pooling layers among successive convolutional layers in the CNN architecture. Therefore, the network can focus only on the important modules generated by convolutional layers.

### (5) Fully-connected Layers

This layer locates at the very end of the entire network and plays the role of a classifier, often accompanied by the soft-max classifier to make the classification decision of the input.

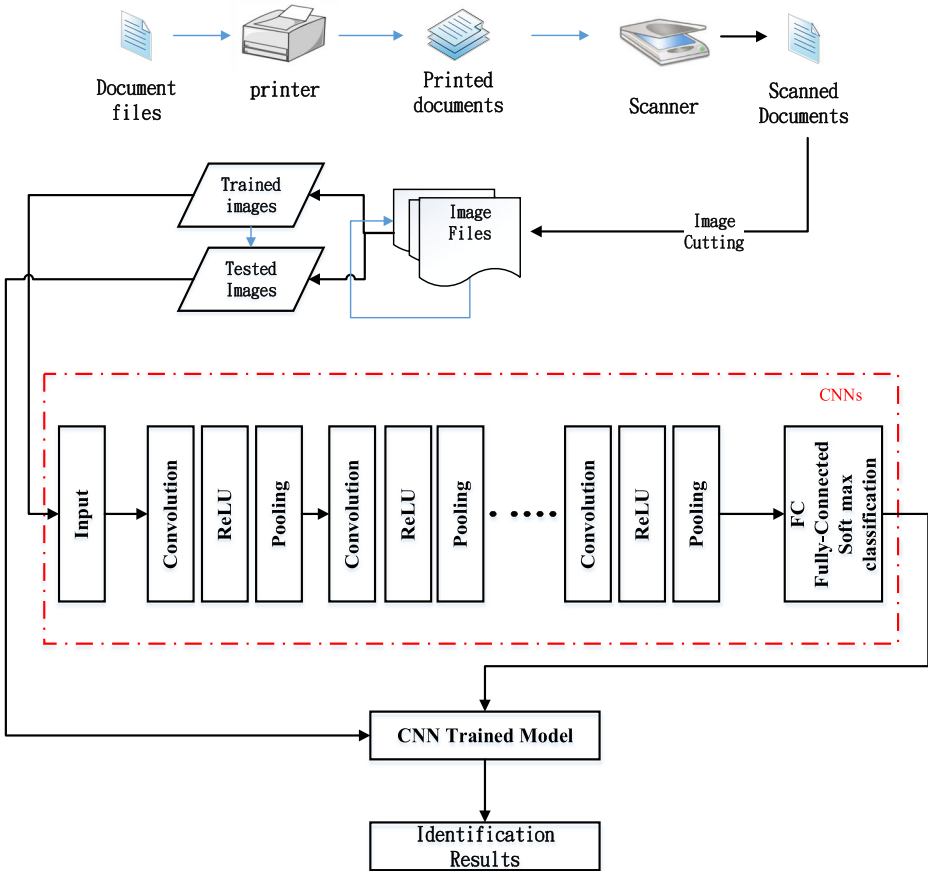
In the convolutional neural network [25], the type of the network layer is arranged according to the different application objectives. The convolution layer, ReLU, the pooling layer forms a core unit. According to the different contents of the input data, the parameters will be adjusted in order to establish a neural network that meets the requirements of the system.

The training of convolutional neural network is mainly operated through the back propagation algorithm and stochastic gradient descent with momentum algorithm. Back propagation can be divided into two phases: propagation and weight update. It technically calculates the gradient of the loss function, and then feedback to the optimization method for weight update [43]. Compared to the standard gradient descent method using the entire dataset, the stochastic gradient descent method for each period of interaction uses the training subset called mini-batch. It minimizes the error function by updating parameters of weight and bias, and the parameter updating is calculated by the gradient of the loss function fed back by the back propagation [43].

## 2.7 Deep learning based classification

The procedure of classification system in this study is shown in Fig. 7 with the following steps:

- (1) Same as the procedure of 3.1(1) printing documents and 3.1(2) digitizing documents.
- (2) Feature extraction and feature selection and classification are all substituted by the convolutional neural network. The different combination of convolutional layer, ReLU layer and pooling layer form a multi-layer neural network. Apply the system to train samples, automatically learn the characteristics of the images from different printers. The fully-connected layer with soft-max operation plays the role of classifier.



**Fig. 7** The flowchart of CNNs system

Based on the above steps, the successful design of the layer architecture of CNNs is the core of the identification system to achieve high accuracy rate. The design of the network layer and the parameters are described as following:

- (1) Input layer: the network layer reads the training samples and generates output to the convolutional layer.
- (2) Convolutional layers: a series of filters are used to convolve the images from the input layer and generate the feature map through convolutional layer. This study will apply three convolutional layers, each of the filter size of convolutional layers is  $5 \times 5$ , zero-padding is set to 2, and stride is set to 1. The first and second convolutional layers use 32 filters to scan the image, and the third convolutional layer uses 64 filters.
- (3) ReLU layers: each convolutional layer is generally followed by the ReLU layer which primarily helps the training to converge quickly and avoids overfitting. The standard non-linear equation  $f(x) = \max(0, x)$  is used in the experiments.
- (4) Pooling layers: a pooling layer reduces the dimensions of the images for subsequent network layers. Each pooling layer has a window dimension of  $2 \times 2$ , and stride is set to 2. It means that the window scans the image and moves two pixels right and down each

time. The first and second pooling layers use the max-pooling and the last one uses average-pooling.

- (5) Fully-connected Layers: only one fully-connected layer is used in the experiments with a soft-max layer and a classification layer to classify the image source of the printer. Since 12 printers are used, the parameter setting of the category is 12.

The rest of the network architecture is set as follows: the initial learning rate of neural network is 0.001. After every 10 cycles of training, the rate will be reduced by 90%. Each experiment will perform 30 cycles of training and each training will process 100 images. Each training sample is randomly selected and each trained image is labelled which will not be selected as testing.

Table 3 summarizes the network layer design of the overall CNNs architecture, where "CONV + POOL<sub>max</sub>" represents the convolutional layer and followed by the use of the maximum generalized pooling layer. The "CONV + POOL<sub>avg</sub>" is followed by the use of pooling layer of average generalization.

### 3 Experiments and discussion

Before we perform the experiments, one issue must be done in advance which we call the pre-tests. The purpose of pre-tests is to determine and verify whether the property of the paper will affect the identification accuracy for printer source identification. It should be conducted before the research flowchart which is shown in Fig. 5.

#### 3.1 Pre-tests

It is first necessary to ensure that the printers are set in black and white mode even color printer is used. As we know, most of the printed documents are printed in black and white. For that reason, we focus on colorless documents. There are types of color printer which is used in this study then we employ the same treatment. To get a fairer outcome, we print in the black-and-white(B&W) printer settings for all the experimental data.

In the pretest experiment, USB microscope is applied to acquire more detailed figure information for text and image documents. As shown in Table 4, there are not significant differences for character "e" between color printer and B&W settings. In contrast, there are

**Table 3** Summary of parameters for CNNs

Layers	1	2	3	4
Type	CONV + POOL <sub>max</sub>	CONV + POOL <sub>max</sub>	CONV + POOL <sub>avg</sub>	FC
Filter numbers	32	32	64	12
Filter size	5 × 5	5 × 5	5 × 5	–
Convolution stride	1 × 1	1 × 1	1 × 1	–
Pooling size	2 × 2	2 × 2	2 × 2	–
Pooling stride	2 × 2	2 × 2	2 × 2	–
Padding size	2 × 2	2 × 2	2 × 2	–

**Table 4** Microscopic images of character “e” and Lena in different printer setting

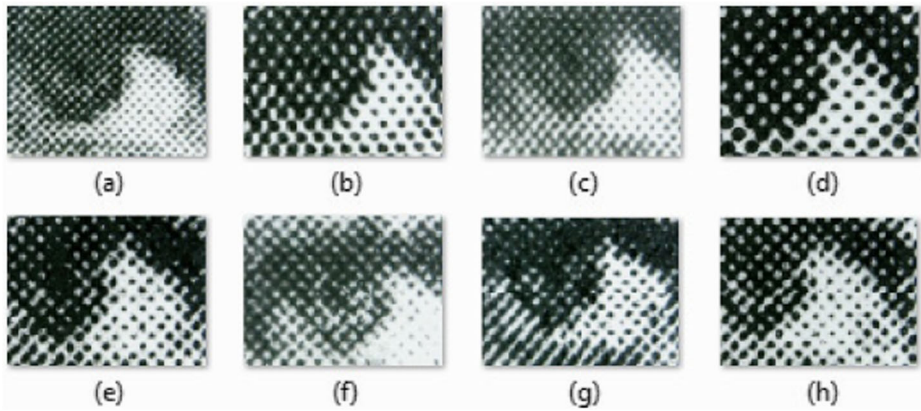
	Character “e” in microscopic image 500x		Lena in grayscale image	
Printer Brand	Color Printing	B&W Printing	Color Printing	B&W Printing
HP LaserJet Pro 300 Color MFP				
HP LaserJet CP1025nw Color				
OKI C5950				

clear differences among image Lena. However, when the images are printed in black printer documents such as are shown in Fig. 8, it is very hard to compare between the color and black.

The pre-tests also analyze different paper quality and the selection of paper material will be examined. During pre-tests, English character “e” will be printed from the same printer HP LaserJet Pro 400 MFP M425dn and then scanned for three different pre-test experiments. For each experiment, 1200 images sample are examined in each character for different type of paper. The images are classified using SVM engine to build the prediction model of the paper sources. There are 500 images which are selected from each paper as training data and another 300 images for test data. The results are tabulated in Table 5 where the individual identification capability from 10 different feature sets of Sec. 2.2 is examined.

The second column of Table 5 lists the first pre-test results while six white paper with the same brands (Paperone 80g) is used. It can be seen in the second column that using the same white paper of the same brand and texture, the average accuracy rates from ten different filters is 32.88%. Generally speaking, the accuracy rate of each filter set is low and it indicates that using identical paper for same printer, it is not possible to do the paper identification.

The third column of Table 5 lists the second pre-test results while the white paper with 6 different textures (six white papers with visually different texture) is applied. Conversely, the average accuracy rate is 87.06% which indicates better classification capability and the accuracy rate of each filter set is significantly higher than the rates of the first pre-test.



**Fig. 8** Microscopic images for Lena in black printer setting. The printer brand and models are (a) Avision AM/MF 3000, (b) HP LaserJet Pro 200 Color P. M251nw, (c) LaserJet Pro 500 MFP M570dn, (d) HP Color LaserJet CP3525, (e) HP LaserJet Pro CP1025 (f) HP LaserJet 4300,, (g) HP LaserJet 4200dtn, (h) HP LaserJet M1132 MFP

The fourth column of Table 5 lists the third pre-test results while the paper with six different color (white, red, green, yellow, pink and blue) is applied. Similarly, the accuracy rate of each filter set is significantly higher than the rates of first pre-test and the average accuracy rate is 86.00%. Therefore, the paper color is also another attribute to affect the printer identification. Due to this fact, the study of printer source identification should be fair without the influence of paper texture or color. Therefore, the same brand of white paper (i.e. Paperone 80 g) is applied in this research for fair comparison purpose. This same setting is also performed in the previous studies [8, 33–36, 53, 55, 56].

### 3.2 Data samples

The study of printed document is categorized into text and image documents. To validate and compare the document, the samples are distinguished not only for text document with different language scripts but also for image document with different samples.

**Table 5** The accuracy rates for character “e” based on different paper texture and color

Filter	Accuracy rates for paper identification		
	Same white paper(%)	Different texture in white paper (%)	Different color paper (%)
LBP	22.58	92.14	79.42
GLCM	33.88	91.24	95.85
DWT	30.03	67.52	77.73
Gaussian	33.06	90.46	93.85
LoG	27.41	84.24	74.91
Unsharp	29.18	88.02	88.81
Wiener	47.32	93.14	75.86
Gabor	30.90	82.14	89.48
Haralick	34.14	90.61	95.30
Fractal	40.34	91.04	88.83
Average(%)	32.88	87.06	86.00

### 3.2.1 Text document

This study examined text document for English character “e”, Chinese character “永”, Arabic character “ع”, and Japanese character “シ”. There are different reasons to select those characters. For example, English character “e” is widely analyzed from various literatures [36, 55]. On the other hand, each stroke of Chinese character “永” represents the basic strokes for Chinese calligraphy [55]. Arabic character “ع” and Japanese character “シ” are selected since both characters are with few number of strokes in its language category. All the scanned text images are converted into grayscale in bitmap file format (BMP). After digitizing the document, the 10 pt. text images are cropped by using software Netbean IDE 8.0 with the pixel size  $51 \times 51$  and the file size is 3.64 kilobytes for each image. As tabulated in Table 6, at least 1200 image samples are examined for each character from one printer source. Hence, the number of sample in each text document is at least 14,400 images from 12 printers.





### 3.2.2 Image document

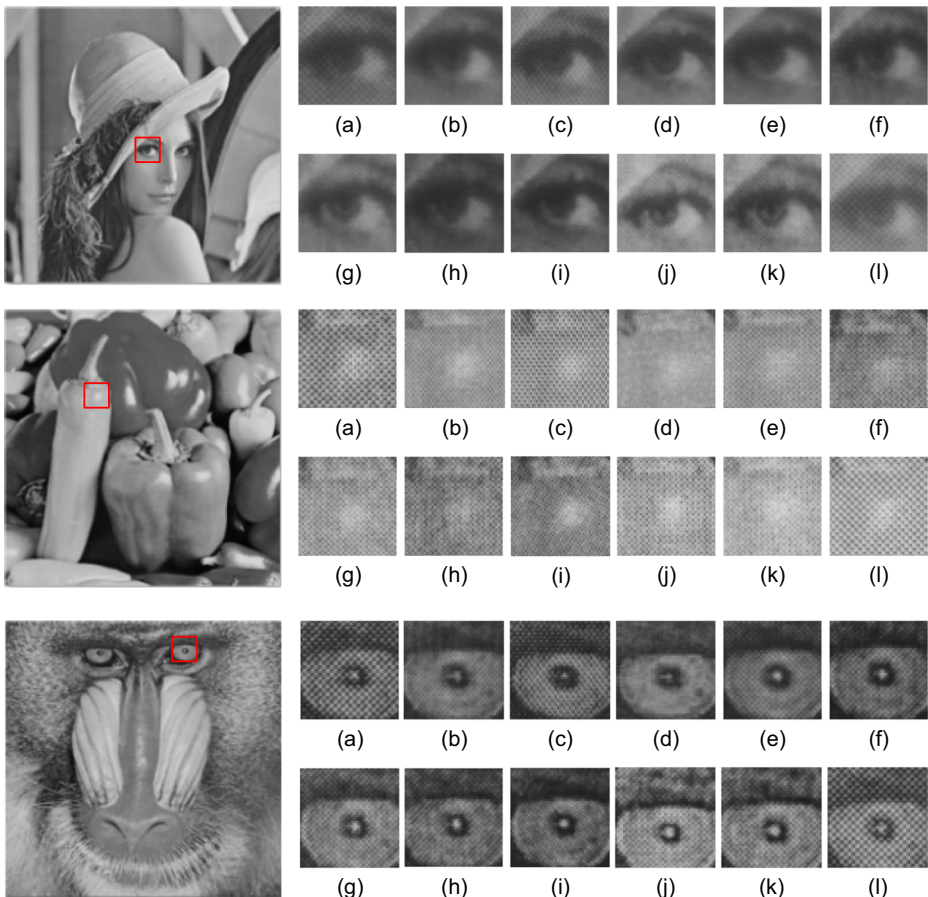
Lena, Peppers, Baboon, and Wikipedia images are evaluated in this study. From the image documents as shown in Figs. 9 and 10, the dimension size of each image patch is also  $51 \times 51$  with 3.64 kilobyte. Each image will be extracted at least 1200 image patch samples from each printer. Thus, the total number of samples for each image document is at least 14,400 images from 12 printers.

## 3.3 Experimental results

In this study, seven experiments are conducted to verify the proposed method as described in Sec. 2.5. We focus on printed source identification for English, Arabic, Chinese and Japanese Character using the combined feature filters and decision fusion model of feature selection. In addition, deep learning based classification system is also performed and compared.

**Table 6** Data information for text documents

	Character Name	Image sample	Dimension	Size	Total data for each printer	Total data proceeded
(1)	“e”, Arial 10 pt		51 x 51	3.64 kb	1,200	14,400
(2)	“永”, Arial 10 pt		51 x 51	3.64 kb	1,200	14,400
(3)	“ع”, Arial 10 pt		51 x 51	3.64 kb	1,200	14,400
(4)	“シ”, Arial 10 pt		51 x 51	3.64 kb	1,200	14,400

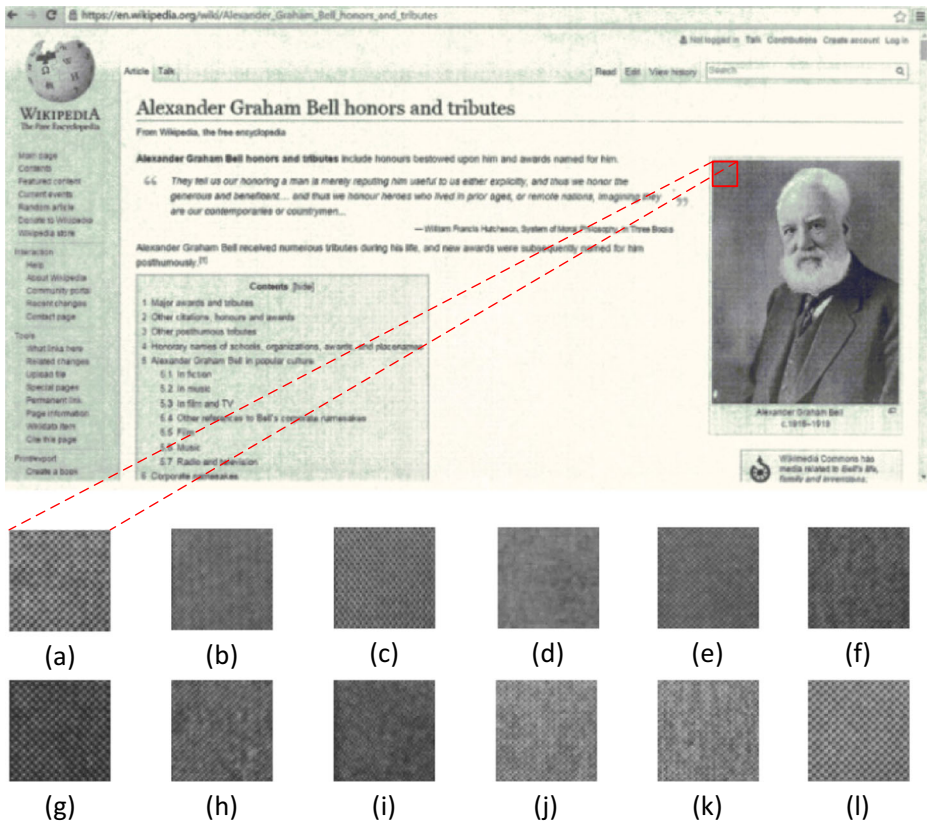


**Fig. 9** Image patch samples for Lena, Peppers, and Baboon from 12 different printers in  $51 \times 51$  pixel size. The printer brand and models are (a) Avision AM/MF 3000, (b) HP LaserJet Pro 200 Color P. M251nw, (c) LaserJet Pro 500 MFP M570dn, (d) HP Color LaserJet CP3525, (e) HP LaserJet Pro CP1025 (f) HP LaserJet 4300, (g) HP LaserJet 4200dtn, (h) HP LaserJet M1132 MFP, (i) HP LaserJet Pro 400 MFP M425dn, (j) HP LaserJet M1522nf, (k) HP LaserJet Pro M1536dnf, and (l) OKI C5900

### 3.3.1 Experiment I: feature selection

**Feature selection techniques** The five feature selection methods are implemented to gain the most computationally effective and efficient features before performing classification. They are Plus-2-Minus-1 (P2M1), Plus-3-Minus-2 (P3M2), Plus-4-Minus-3 (P4M3), Sequential Forward Floating Search (SFFS), and Sequential Backward Floating Selection (SBFS) and the algorithms are written in Java. Regarding about how many features should be selected and how much time is needed, these features are grouped into 10 major groups such as GLCM (22 features), DWT (12 features), Gaussian (21 features), LoG (21 features), Unsharp (21 features), Wiener (64 features), Gabor (48 features), Haralick (14 features), fractal (24 features) and LBP (59 features). In addition, all features are selected from 1200 images in each letter and image sources.





**Fig. 10** Image patch samples for “Wikipedia.org” from 12 different printers in  $51 \times 51$  pixel size. The printer brand and models are (a) Avision AM/MF 3000, (b) HP LaserJet Pro 200 Color P. M251nw, (c) LaserJet Pro 500 MFP M570dn, (d) HP Color LaserJet CP3525, (e) HP LaserJet Pro CP1025 (f) HP LaserJet 4300, (g) HP LaserJet 4200dn, (h) HP LaserJet M1132 MFP, (i) HP LaserJet Pro 400 MFP M425dn, (j) HP LaserJet M1522 nf, (k) HP LaserJet Pro M1536dnf, and (l) OKI C5900

In this experiment, Arabic character “ﻻ” in Arial typeface with the font size 10 point is evaluated for feature selection. The SVM model setting is adopted from Mikkilineni [34] and Tsai [55] who applied 500 images for training and 300 images for testing. The feature selection algorithm is executed by adding or removing a feature one at a time to find the optimum identification rate. The selection order during execution will be recorded to choose the most important features. Each feature selection algorithm is repeated 10 times for different data sets and the diagram of accuracy rate versus number of features is plotted to decide the cut off number for the most important features. Table 7 lists the differences among accuracy prediction based on the number of features for GLCM feature set using P2M1 selection algorithm. It can be clearly seen that the number of feature considerably affect the accuracy prediction. The less number of features could lead to the low accuracy predictions. As shown in Table 7, two features are selected which could attain only 72.19% for accuracy prediction and seven features are selected which could attain 95.44% for accuracy prediction. However, increasing number of features does not always significantly improve the prediction accuracy. It can be seen that the accuracy prediction is topped at 16 features selected from Table 7 with the

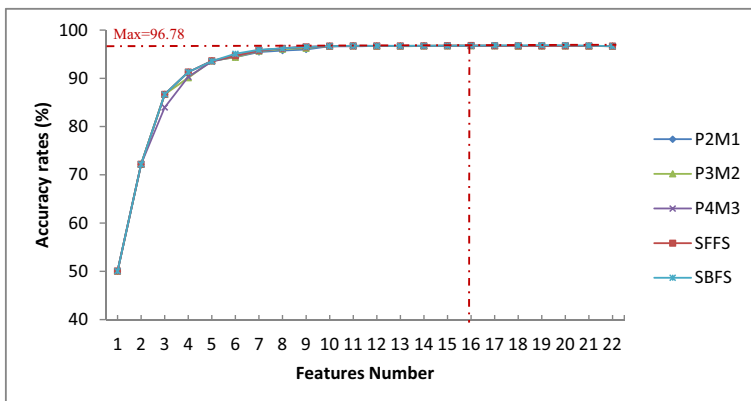
**Table 7** GLCM feature set based on P2M1 selection algorithm

Number of Features	Selected Feature Sets in Order	Accuracy Prediction(%)
1	20	50.06
2	20,13	72.19
3	20,13,2	86.67
4	20,13,2,18	91.28
5	20,2,18,14,16	93.61
6	20,2,18,14,16,15	94.39
7	20,2,18,14,16,5,13	95.44
8	20,2,18,14,16,5,13,17	95.75
9	20,2,18,14,16,13,17,11,8	95.97
10	20,2,18,16,13,17,11,8,12,10	96.67
11	20,2,18,16,13,17,11,8,12,10,3	96.72
12	20,2,18,16,13,17,11,8,12,10,3,19	96.69
13	20,2,18,16,13,17,11,8,12,10,3,5,4	96.69
14	20,2,16,13,17,11,8,12,10,3,5,4,6,14	96.72
15	20,2,16,13,17,11,8,12,10,3,5,4,6,14,18	96.75
16	20,2,16,13,17,11,8,12,10,3,4,6,14,18,19,22	96.78
17	20,2,16,13,17,11,8,12,10,3,6,14,18,19,22,5,7	96.75
18	20,2,16,13,17,11,8,12,10,3,6,14,18,19,22,5,7,4	96.72
19	20,16,13,17,11,8,12,10,3,6,14,18,19,22,5,7,4,15,1	96.67
20	20,16,13,17,11,8,12,10,3,6,18,19,22,5,7,4,15,1,2,9	96.72
21	20,16,13,18,11,8,12,10,3,6,18,19,22,5,7,4,15,1,2,9,21	96.69
22	20,16,13,19,11,8,12,10,3,6,18,19,22,5,7,4,15,1,2,9,21,14	96.67

maximum accuracy prediction is 96.78%. This experiment suggests that the wise feature selection from each feature set can achieve the best accuracy rate with less number of features.

**Deciding the most important  $\lambda$  features** In the context of obtaining the order list after feature selection, it takes the decision fusion of 5 feature selection methods for aggregation which can increase the model accuracy, decrease the overall computation time and reduce overfitting with less redundant data for decision.

The diagram of accuracy rate versus number of features is plotted to decide the cut off number for the most important features. As Fig. 11 shows, the accuracy rates for different



**Fig. 11** The most important  $\lambda$  features based on accuracy rates of 5 feature selection

feature sets increase considerably from feature number 0 to feature number 5 and is topped at near 16 features. Hence, the cut off number is set to 16. Thus, the method of feature selection fusion from P2M1, P3M2, P4M3, SFFS, and SBFS achieved 16 optimum features from 22 features for GLCM.

According to the optimum features that have already been recognized, we further decide which features should be selected from the feature order list. Afterwards, they are numerically sorted in ascending order based on the frequency times. Hereafter, as shown in Table 8, we select 16 best indexed features to get optimum feature sets for GLCM. Based on the convex property of the plot, the most important  $\lambda$  features can be decided from the highest rate according to the plot. Using the recorded feature-selection order, the counter-based decision fusion algorithm is utilized to decide the final top 16 selected features. The adaptive decision theoretical model of feature selection is implemented in this study in order to reduce the total evaluation time without the loss of accuracy while the most important  $\lambda$  features are selected. The number of chosen features is determined based on the accuracy rate for all 306 features.

**Table 8** Selected features for each feature set

Group name	Number of features	Number of Selected Feature	Feature order with the numbered features
GLCM	22	16	2,3,4,5,6,8,10,11,12,13,14,16,17,18,19,20
DWT	12	12	23,24,25,26,27,28,29,30,31,32,33,34
Gaussian	21	15	35,37,38,42,44,45,46,47,48,49,50,52,53,54,55
LoG	21	18	56,57,58,59,60,61,62,65,66,68,69,70,71,72,73,74,75,76
Unsharp	21	14	77,78,79,82,86,87,88,89,90,92,94,95,96,97
Wiener	64	30	106,108,109,110,111,113,122,124,125, 126,127,128,131,133,135,138,140,141,142,143, 144,145,154,155,156,157,158,159,160,161
Gabor	48	27	163,165,168,169,170,171,175,176,177,178,181, 182,183,184,185,187,188,193,195,199,200,201, 202,203,205,208,209
Haralick	14	12	214,215,216,217,218,219,220,221,222,223
Fractal	24	23	224,225,226,227,228,229,230,231,232,233,234, 235,236,237,238,240,241,242,243,244,245,246,247
LBP	59	57	248,249,250,251,252,254,255,256,257,258,259,260, 261,262,263,264,265,266,267,268,269,270,271,272, 273,274,276,277,278,279,280,281,282,283,284,285, 286,287,288,289,290,291,292,293,294,295,296,297, 298,299,300,301,302,303,304,305,306
All	306	222	2,3,4,5,6,8,10,11,12,13,14,16,17,18,19,20,23,24,25,26, 27,28,29,30,31,32,33,34,35,37,38,42,44,45,46,47,48,49, 50,52,53,54,55,56,57,58,59,60,61,62,65,66,68,69,70,71, 72,73,74,75,76,77,78,79,82,86,87,88,89,90,92,94,95,96, 97,106,108,109,110,111,113,122,124,125,126,127,128, 131,133,135,138,140,141,142,143,144,145,154,155,156, 157,158,159,160,161,163,165,168,169,170,171,175,176, 177,178,181,182,183,184,185,187,188,193,195,199,200, 201,202,203,205,208,209,214,215,216,217,218,219,220, 221,222,223,224,225,226,227,228,229,230,231,232,233, 234,235,236,237,238,240,241,242,243,244,245,246,247, 248,249,250,251,252,254,255,256,257,258,259,260,261, 262,263,264,265,266,267,268,269,270,271,272,273,274, 276,277,278,279,280,281,282,283,284,285,286,287,288, 289,290,291,292,293,294,295,296,297,298,299,300,301, 302,303,304,305,306

Same approaches are performed not only for GLCM but also for other nine feature sets and we only list the feature value of  $\lambda$  after the experimental analysis due to the limited space. The selection order during execution is recorded for the most important features in group fusion format as shown in Table 7. After the feature selection, total 306 features can be efficiently reduced to 222 features in Table 8. After the most important features have been decided, the printer source identification can be finally investigated.

### 3.3.2 Experiment II: applying feature filter sets without selection

**Text document** To evaluate the printer source by using the proposed approach, the accuracy rate based on character “e” is first investigated from previous research [8, 33–36]. In this study, we examine English character “e”, Chinese character “永”, Arabic character “ﺍ”, and Japanese character “シ”. Before classification, all images are firstly extracted by using 10 feature extraction sets such as mentioned in Section 2. Then, whole extracted images with numeric characteristic value files are exported to MySQL database and classified based on filter sets by using SVM in the Eclipse of Java environment. Through the help of the tool “grid.py” from [15], the optimum penalty and gamma settings of SVM for the proposed method are found to be 32 and 1, and radial basis function (RBF) kernel is used for SVM classification.

The steps of proposed approaches are listed below:

- (1) 10 sets of images from text document database of 12 printer sources are randomly generated. In each set, there are 500 images which are selected from each printer as training data and another 300 images for testing data. 10 sets of feature filters are then applied for characteristic extraction.
- (2) Apply the SVM engine to build the prediction models using 10 sets of feature filter.
- (3) Feed the test image subsets to the corresponding model trained in step 2 for the printer source prediction.
- (4) Repeat step 1 through 3 ten times to obtain the predicted results.

Table 9 tabulates the accuracy prediction of ten set of filters for text documents. The accuracy prediction rate for character “e”, “ﺍ”, “シ”, and “永” have stable percentages on each

**Table 9** The accuracy prediction among different data sets for English, Arabic, Chinese, and Japanese characters

Filter group	Number of features	English e (%)	Arabic ﺍ (%)	Chinese 永 (%)	Japanese シ (%)	Average (%)
GLCM	22	93.76	96.38	95.56	90.91	94.15
DWT	12	83.52	91.54	85.62	90.74	87.86
Gaussian	21	93.37	96.65	94.87	90.52	93.85
LoG	21	84.57	82.20	81.99	76.13	81.22
Unsharp	21	90.95	90.48	89.46	86.79	89.42
Wiener	64	95.74	96.55	95.19	88.54	94.01
Gabor	48	95.49	95.36	95.67	96.15	95.67
Haralick	14	88.30	94.98	93.50	88.69	91.37
Fractal	24	89.58	90.12	88.24	84.81	88.19
LBP	59	77.66	55.62	58.34	50.23	60.46
All	306	96.16	98.77	98.46	96.11	97.38

feature set of filters. The results of the classification using SVM based on feature extraction without selection. It shows the average prediction for all features and the prediction rate from Gabor filter set is 95.67% which is the highest result among the feature sets. Generally speaking, the accuracy rates of other four feature filter sets (GLCM, Gaussian, Wiener and Haralick) in average are above 90%. In addition, the other feature filter sets which has accuracy rates above 80% is DWT, LoG, and Fractal. Conversely, using LBP features can get the average accuracy rate at 60.46%. It is the lowest accuracy rate to identify the printed sources. Overall, the Gabor features are superior to analyze text document for different alphabet if feature selection is not applied.

**Image document** The steps to process the image document are similar to the operations of text document where the results from the ten filters are tabulated in Table 10. The average accuracy rates for all feature sets are above 90%. It can be seen that Gaussian feature set has the lowest accuracy prediction (94.18%) for image documents. The LBP filters constantly have the highest percentage compared with other filters. The data table shows the average prediction for LBP filter is above 99%. The average percentage of prediction results by using GLCM, LoG, Un-sharp, Wiener, Gabor, fractal, Haralick and LBP are above 95%. Conversely, the classification result for both feature sets (DWT and Gaussian) is under 95%. Among those feature sets, LBP feature sets achieve the highest accuracy ratios with the average percentage at nearly 100% and the Gabor filter achieves the second.

### 3.3.3 Experiment III: using decision fusion

In this section, the feature selection is adopted for text and image data while there are total 306 features and the most important 222 features are selected. The experimental results are tabulated in Table 11. The procedures of this experiment are similar to Experiment II and briefly described below:

- (1) 10 sets of images from image database of 12 printer sources are randomly generated for text and image documents. In each set, there are 500 images which are selected from each printer as training data and another 300 images for test data. The 10 set filters (GLCM,

**Table 10** The accuracy prediction rates among different data sets for Lena, Peppers, Baboon and Wikipedia

Filter group	Number of features	Lena (%)	Peppers (%)	Baboon (%)	Wikipedia of Fig. 7 (%)	Average (%)
GLCM	22	97.50	99.72	97.30	99.10	98.41
DWT	12	96.43	95.82	89.60	95.19	94.26
Gaussian	21	96.56	99.34	83.56	97.26	94.18
LoG	21	94.70	97.37	95.87	97.21	96.29
Unsharp	21	95.48	97.72	94.57	97.37	96.29
Wiener	64	97.40	99.54	98.69	99.62	98.81
Gabor	48	99.25	99.93	99.59	99.92	99.67
Haralick	14	94.77	99.13	93.41	96.89	96.05
Fractal	24	96.78	97.91	97.87	97.02	97.40
LBP	59	99.98	100	99.99	100	99.99
All	306	99.84	99.89	99.89	99.90	99.88

**Table 11** The accuracy prediction rates using feature fusion and selection for text and image documents

Character	Text document		Image	Image document	
	All 306 features (%)	Selected 222 features (%)		All 306 features(%)	Selected 222 features (%)
e	96.16	97.96	Lena	99.84	99.93
↗	98.77	99.43	Peppers	99.89	99.97
永	98.46	99.23	Baboon	99.89	99.98
シ	96.11	97.50	Wikipedia	99.90	99.94
<b>Average</b>	<b>97.38</b>	<b>98.53</b>	<b>Average</b>	<b>99.88</b>	<b>99.96</b>

DWT, Gaussian, LoG, Unsharp, Wiener, Gabor, Haralick, Fractal features and LBP features) are then calculated.

- (2) The feature selection algorithm is executed by adding or removing one feature at a time to find the optimum identification rate. The selection order during execution is recorded to choose the most important features and applying the SVM engine to build the prediction models using 306 features from each filter sets group.
- (3) Using the recorded feature-selection order, the counter-based decision fusion algorithm is used to decide the final top 222 selected features from the results of 10 tests
- (4) Feed the test image subsets to the corresponding model trained in step 3 for the printer source prediction.
- (5) Repeat step 1 through 3 up to ten times to obtain the predicted results

### 3.3.4 Experiment IV: prediction comparison using different scanner resolution

Dot per inch (DPI) is used to describe the resolution number of dots per inch in a digital print and the printing resolution of a hardcopy print. In general, printers with higher DPI can produce clearer and more detailed output. On the other hand, the printed document will be scanned to form the image data where the information content of the final image is determined by the parameter “dots per inch” (dpi) of the optical flatbed scanner. Commercial flatbed scanners allow a wide range of resolution (even thousands of dpi) while higher resolution setting will generate larger file size. To make reasonable comparison for general applications, the images using scanner resolution of 300 dpi and 600 dpi are studied in this experiment. The confusion matrixes of identification for English character “e” results are shown in Tables 12, 13, 14, and 15 respectively. Apparently, the average identification accuracy rate at 600 dpi (96.16% from Table 14) is higher than the one at 300 dpi (92.09% from Table 11) by all features (306 features).

After adopting the feature selection to select the most important features, the average accuracy rate by using 222 features (95.45% from Table 13) is better than the one using all features (92.09% from Table 12) while the scanner resolution is 300 dpi. Under similar situation, the average accuracy rate by 222 features (97.96% from Table 15) is better than the one using all features (96.16% from Table 14) at 600dpi. Perceivably, higher resolution of DPI values generally results higher accuracy rates due to more detailed information is involved from the scanned images and the feature selection can even achieve better identification precision.

**Table 12** The confusion matrix of identification results using 306 features for “e” at 300 DPI

Avg		Predicted (%)											
92.09		1	2	3	4	5	6	7	8	9	10	11	12
Actual	1	99.90	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.00
	2	5.87	92.53	0.00	0.00	0.40	1.20	0.00	0.00	0.00	0.00	0.00	0.00
	3	2.37	0.00	95.20	0.00	0.00	0.00	0.00	0.03	2.30	0.00	0.00	0.10
	4	1.70	0.00	0.13	92.23	0.00	0.00	0.03	1.53	3.40	0.30	0.00	0.67
	5	4.07	0.40	0.10	0.00	94.27	0.90	0.00	0.00	0.27	0.00	0.00	0.00
	6	2.63	0.07	0.00	0.00	0.40	96.90	0.00	0.00	0.00	0.00	0.00	0.00
	7	0.33	0.00	0.00	0.00	0.00	0.00	94.57	0.80	2.13	0.50	1.47	0.20
	8	0.10	0.00	0.37	0.03	0.00	0.00	2.97	74.00	20.17	1.60	0.03	0.73
	9	0.20	0.00	0.33	0.00	0.00	0.00	0.80	4.37	94.30	0.00	0.00	0.00
	10	0.53	0.00	0.00	0.20	0.00	0.00	6.33	3.17	1.47	84.40	3.50	0.40
	11	1.17	0.00	0.00	0.00	0.00	0.00	2.90	0.07	0.00	0.53	95.33	0.00
	12	1.37	0.00	0.00	0.27	0.00	0.00	3.20	0.70	2.27	0.77	0.00	91.43

3.3.5 Experiment V: effect of font type and text size

The effect of different font type and text size is also investigated. There are two variation of shape (type) and font size, the font type Arial with the size of 8, 10, 12, and 14 point and Courier new with the size of 8, 10, 12, and 14 point for character “≈” by using all 306 features and 222 selected features. Table 16 tabulates the accuracy rates of the classifiers on different

**Table 13** The confusion matrix of identification results using 222 features for “e” at 300 DPI

Avg		Predicted (%)											
95.45		1	2	3	4	5	6	7	8	9	10	11	12
Actual	1	99.93	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.03	0.00
	2	0.10	98.60	0.00	0.00	0.37	0.93	0.00	0.00	0.00	0.00	0.00	0.00
	3	0.57	0.00	98.13	0.00	0.00	0.00	0.00	0.00	1.30	0.00	0.00	0.00
	4	0.57	0.00	0.27	95.60	0.00	0.00	0.00	0.77	2.10	0.13	0.00	0.57
	5	0.47	0.53	0.00	0.00	98.63	0.20	0.00	0.00	0.17	0.00	0.00	0.00
	6	0.13	0.23	0.00	0.00	1.10	98.53	0.00	0.00	0.00	0.00	0.00	0.00
	7	0.00	0.00	0.00	0.00	0.00	0.00	96.30	0.47	1.60	0.23	1.00	0.40
	8	0.00	0.00	0.30	0.00	0.00	0.00	2.63	80.00	15.70	0.83	0.03	0.50
	9	0.00	0.00	0.30	0.00	0.00	0.00	0.60	4.63	94.40	0.00	0.00	0.07
	10	0.00	0.00	0.00	0.10	0.00	0.00	3.30	1.73	0.77	91.53	2.13	0.43
	11	0.23	0.00	0.00	0.00	0.00	0.00	1.73	0.00	0.00	0.70	97.23	0.10
	12	0.03	0.00	0.00	0.13	0.00	0.00	1.37	0.67	0.57	0.70	0.00	96.53

**Table 14** The confusion matrix of identification results using 306 features for “e” at 600 DPI

Avg		Predicted (%)											
		1	2	3	4	5	6	7	8	9	10	11	12
Actual	1	99.70	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.30	0.00
	2	0.00	96.37	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	3.60	0.00
	3	0.13	0.00	97.80	0.07	0.00	0.00	0.00	0.00	0.53	0.00	1.47	0.00
	4	0.10	0.00	1.23	94.33	0.00	0.00	0.00	0.13	1.07	0.17	1.73	1.23
	5	0.33	0.03	0.00	0.00	98.83	0.00	0.00	0.00	0.03	0.00	0.77	0.00
	6	0.17	0.00	0.00	0.00	0.00	98.93	0.00	0.00	0.00	0.00	0.90	0.00
	7	0.00	0.00	0.00	0.00	0.00	0.00	94.70	0.90	0.87	0.33	3.20	0.00
	8	0.00	0.00	0.00	0.00	0.00	0.00	0.07	88.83	8.63	0.27	2.03	0.17
	9	0.03	0.00	0.03	0.00	0.00	0.00	0.20	4.23	93.90	0.07	1.50	0.03
	10	0.00	0.00	0.00	0.00	0.00	0.00	1.17	0.93	0.93	94.03	2.77	0.17
	11	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.00	0.00	0.67	99.13	0.00
	12	0.00	0.00	0.00	0.13	0.00	0.00	0.07	0.17	0.50	0.47	1.37	97.30

shape and dimension. For different scenarios, the accuracy rates generally increase while the font size increases as shown in Table 16. For example, the accuracy rates for classifying 14 point text documents are better than those for 8 point text documents. Not only the image dimension increases for font size but also the accuracy rates increase due to more detailed information included.

**Table 15** The confusion matrix of identification results using 222 features for “e” at 600 DPI

Avg		Predicted (%)											
		1	2	3	4	5	6	7	8	9	10	11	12
Actual	1	99.83	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.00
	2	0.00	98.57	0.00	0.00	0.20	0.00	0.00	0.00	0.00	0.00	1.23	0.00
	3	0.07	0.00	99.20	0.13	0.00	0.00	0.00	0.00	0.07	0.00	0.53	0.00
	4	0.03	0.00	0.77	97.03	0.00	0.00	0.00	0.07	0.73	0.07	0.60	0.70
	5	0.00	0.00	0.00	0.00	99.77	0.00	0.00	0.07	0.00	0.00	0.17	0.00
	6	0.03	0.00	0.00	0.00	0.00	99.83	0.00	0.00	0.00	0.00	0.13	0.00
	7	0.00	0.00	0.00	0.00	0.00	0.00	98.00	0.60	0.70	0.27	0.43	0.00
	8	0.00	0.00	0.00	0.00	0.00	0.00	0.23	92.40	6.30	0.10	0.50	0.47
	9	0.00	0.00	0.00	0.00	0.00	0.00	0.13	4.03	95.60	0.03	0.17	0.03
	10	0.00	0.00	0.00	0.00	0.00	0.00	0.60	0.57	0.63	96.63	1.30	0.27
	11	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.00	0.00	0.33	99.47	0.00
	12	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.27	0.33	0.13	99.17



**Table 16** The average accuracy prediction rates based on different font type and size for character “ا”

Font type	Dimension	Size	306 features (%)	222 features (%)
Arial 8 pt	43 × 43	2.90 kb	96.66	97.98
Arial 10 pt	51 × 51	3.64 kb	98.77	99.43
Arial 12 pt	59 × 59	4.50 kb	98.67	99.34
Arial 14 pt	67 × 67	5.50 kb	98.79	99.36
Courier New 8 pt	43 × 43	2.90 kb	97.93	98.71
Courier New 10 pt	51 × 51	3.64 kb	98.58	99.08
Courier New 12 pt	59 × 59	4.50 kb	98.53	99.38
Courier New 14 pt	67 × 67	5.50 kb	98.63	99.35
Average			98.32	99.08

Even the rate for Arial 12 (98.67%) is slightly lower than the one for Arial 10 point (98.77%) from Table 16 using 306 features and 222 features, the results are very close and compatible. Accordingly, the average of accuracy rates using 306 features is 98.32% and the one applying 222 features is 99.08%. As expected, the implementation of decision fusion model for feature selection can effectively achieve higher classification outcomes than the one using all features.

### 3.3.6 Experiment VI: comparison of different CNNs network architecture

In order to compare the deep learning based classification system, the network structure will be designed according to the amount of data and the complexity of the image content. It is not necessary that more hidden layers included will accomplish higher accuracy results due to the overfitting in the network. The design principle is not only to get a concise network architecture for less training time but also to achieve high accuracy rate.

In this experiment, the network architecture has three different depths, either using 1, 2 or 3 convolutional layers respectively, and each convolutional layer will be equipped with ReLU layer and pooling layer. Therefore, total 7 layer, 10 layer and 13 layer neural network models are designed for comparison. The text image inputs include characters for Chinese, Japanese, Arabic, and images including Lena, Peppers and Baboon. There are 500 training samples from the images in each category for training and randomly selected 300 samples from the remaining images as test samples. Each text and image of the experimental samples are sent to the 7, 10 and 13 layer CNNs network (details in Table 3) for training, classified respectively and the results are shown in Tables 17, 18, and 19.

According to Table 17, the highest accuracy rate is 2 conv CNNs for Chinese character “的” at 98.29% and the lowest one is 88.81% for “文” at 1 conv CNNs. If increasing the convolutional level for “文”, the accuracy rate can be raised up to 92.38%. In Table 18, the highest accuracy rate for Japanese character is “あ” at 99.44% and the lowest one is 94.08% for “ノ” for 1 conv CNNs. According to Table 19, the highest recognition rate for Arabic character is 99.22% of “” and the lowest one is “95.08%” for “پ” for 1 conv CNNs. From above

**Table 17** The accuracy rates for Chinese characters

1 conv (7 Layers)	96.75%	94.86%	97.67%	88.81%
2 conv (10 Layers)	98.06%	97.90%	98.29%	92.29%
3 conv (13 Layers)	95.28%	96.48%	97.67%	92.38%

**Table 18** The accuracy rates for Japanese characters

	ア	あ	シ	し	ノ	の
1 conv (7 Layers)	97.86%	99.44%	96.08%	95.11%	94.08%	97.44%
2 conv (10 Layers)	96.94%	99.16%	95.47%	93.00%	91.61%	96.52%
3 conv (13 Layers)	94.72%	97.66%	93.08%	89.8%	87.91%	94.80%

observation, the amount of strokes in text has significant influence during the classification and the character with less strokes is much harder for classification even convolutional level increased. However, the accuracy rates for different CNNs systems are extremely high for image Lena, Peppers and Baboon in Table 20 since more complicated content of the image provides more useful information for the neural network.

From the experimental data in Tables 17, 18, 19, and 20, it is clear that 7 or 10 layers CNNs can achieve the highest accuracy rate for most of the text data. Since natural images contain more complex texture than character text, it is understandable that more level of layer structure in CNNs could extract better abstraction from the feature map. Therefore, 10 layer CNNs achieves the best accuracy in Table 20. Consequently, this comparison suggests that different structure of CNNs should be applied based on the complexity of input data.

### 3.3.7 Experiment VII: comparison of SVM based classification system and CNNs based classification system for scanned images

It is an interesting topic to discover the performance behavior between the SVM based classification system and CNNs based classification system. Table 21 tabulates the printer source classification for Chinese character “永”, Japanese character “シ”, Arabic character “ﺍ”, Lena, Peppers and Baboon images respectively. Since the feature based SVM system is optimized for Arabic character “ﺍ”, there is no doubt that the accuracy rates for “ﺍ” are the best among characters. Deep learning based system also achieves high accuracy rates for text data and the difference is less than 1% in average for character “永”, and “シ”.

On the other hand, the SVM based system performs very well for Lena, Peppers and Baboons, and the results for both systems are comparable. In most cases, 1 conv CNNs structure based deep learning system can achieve better classification results for images than text documents while the input is from scanned images.

## 3.4 Discussion

### 3.4.1 Performance evaluation

Currently, there are no standard benchmark tests of digital forensics for printer source identification. To make fair comparison, the same experiments are performed for the proposed

**Table 19** The accuracy rates for Arabic characters

	ا	ﺍ	ﺏ	ﺓ	ﻉ	ﻊ
1 conv (7 Layers)	95.08%	98.44%	98.47%	99.22%	98.72%	98.39%
2 conv (10 Layers)	94.00%	98.72%	98.89%	98.50%	97.97%	97.92%
3 conv (13 Layers)	82.58%	96.36%	96.61%	94.61%	97.06%	87.42%

**Table 20** The accuracy rates for Images

	Lena	Peppers	Baboon
1 conv (7 Layers)	99.93%	99.91%	99.93%
2 conv (10 Layers)	99.97%	99.91%	99.96%
3 conv (13 Layers)	99.65%	99.75%	99.79%

approach and the techniques from literatures and the results are shown in Figs. 12 and 13. From Fig. 12, the accuracy rate of character “e” using our proposed method is 97.96% which is better than the one from Mikkilineni’s algorithm [35], while Ferreira [8] achieved 97.60% that is also compatible. The experiment results for text document of character “永” and “シ” are also superior to the previous studies of [55, 56].

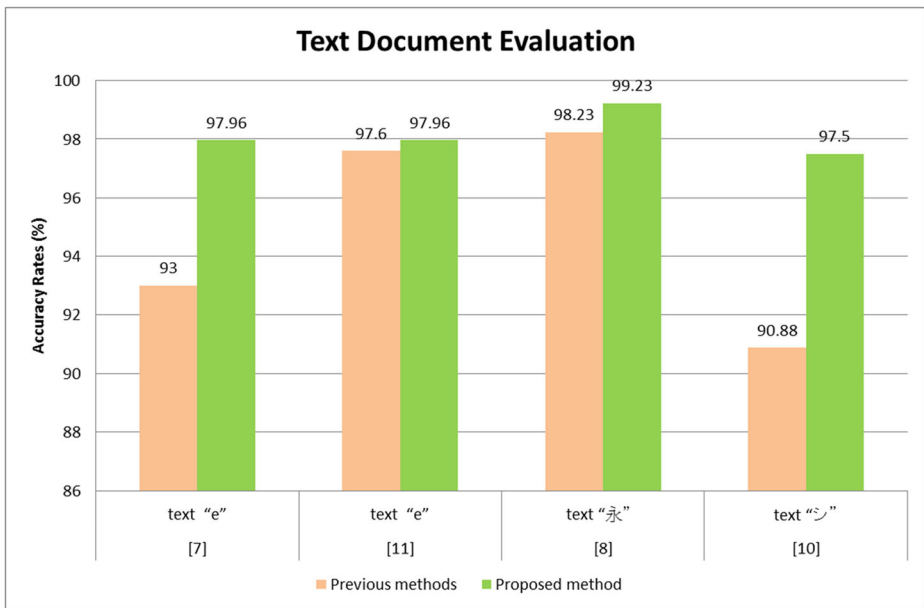
As we can see from Fig. 13, Ferreira [8] analyzed frame of image document of Wikipedia in printer sources identification and attained 98.47%. Choi et al. [5] used GLCM and DWT with the 384 statistical features to extract image documents (Baboon, Lena, Peppers, etc.) for identifying color laser printer. There are three kinds of experiments performed in [5]: the brand identification test, the toner identification test and the model identification test. The brand identification tests whether the algorithm can make a difference among different color laser printer brands or vendors. In the toner identification test, whether the difference of the color toners would make the difference in printed images or not is performed. The model identification test checks whether the algorithm can discern among different color laser printer models. In their experiments, 8 color laser printers from 4 different brands are used and they achieved 88.75% accuracy for the model identification. In addition, Kim and Lee [23] also identified color laser printer to analyze halftone images in the discrete Fourier transform (DFT) domain where they got 94.4% accuracy rate. As shown in Fig. 13, the average accuracy rate of our proposed method is 99.96% (using Lena, Pepper, Baboon, and Wikipedia images) which is superior to previous studies in image document evaluation. It demonstrates that our proposed approach can achieve the best identification rates and the technique can effectively identify the printer sources based on either text or image documents.

### 3.4.2 Paper quality

In Section 3.1, this study performs the pre-test which examines the influence of the paper and the results conclude the use of different paper will affect the classification results. Different types of paper have varied surface and structure which can be magnified by micro-scale imaging for observation purpose [51]. The choice of paper can also influence the quality of image and printed documents [49]. A printer will produce different image contingent on the

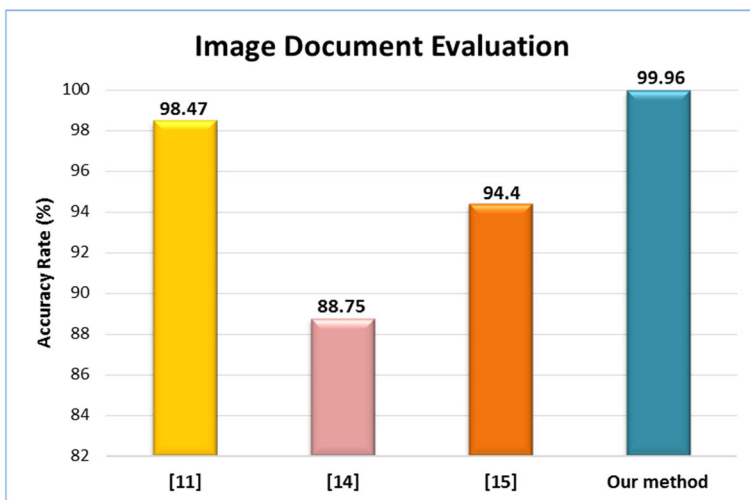
**Table 21** Comparison of feature based SVM classification system and CNNs based classification system for scanned images

	シ	⇨	Lena	Pepper	Baboon
SVM(306 features)	98.46%	96.11%	98.77%	99.84%	99.89%
SVM(222 features)	99.23%	97.50%	99.43%	99.93%	99.97%
1 conv (7 Layers)	96.75%	97.88%	98.47%	99.93%	99.93%
2 conv (10 Layers)	98.06%	94.64%	98.89%	99.97%	99.91%
3 conv (13 Layers)	95.28%	72.32%	96.61%	99.65%	99.75%



**Fig. 12** Evaluation diagram of accuracy prediction rates for text document

orientation of the page because the paper surface may not be perfectly flat. Additionally, the surface roughness of coated and uncoated papers influence the occurrence of print mottles [20]. It means the absorption of ink and fountain solution cause non-uniformity of ink transfer because of the surface roughness of paper [46]. The selection of paper will have strong impact for printing properties including a huge variety of characteristics such as its particle, chemical composition, and shape [19]. Accordingly, to avoid the influence of paper texture or color, the same texture and color paper is used in printed document source identification which policy is also adopted in previous study [8, 35, 55].



**Fig. 13** Evaluation diagram of accuracy prediction rates for image document

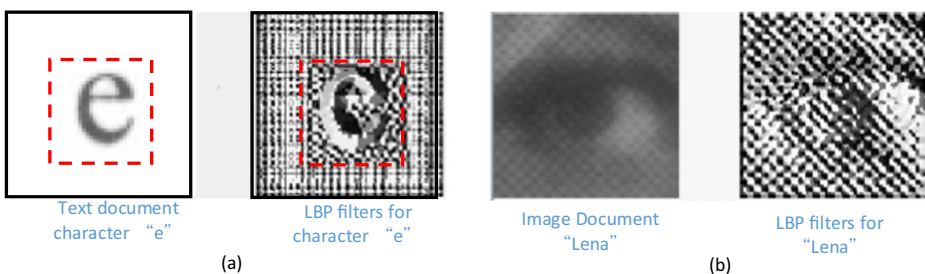
### 3.4.3 ROI consideration

In Section 1, this research discusses several previous studies regarding how to implement the intrinsic signatures for printer source identification. Since Toner particles are specifically melted by the heat of the fuser, and are thus bonded to the paper. The electromechanical imperfections in printing process will cause the toner variations in the development stage. Such intrinsic texture signatures from the printer stuck on a paper document including shapes, sizes and patterns authentically provide the substantial information as a guide for researchers to distinguish and classify the printer sources. The information outside the printing area generally contains only paper texture information or non-uniform motor drifting trace mark during the printing process. Generally speaking, such signals are irrelevant and can be removed during the feature extraction.

Figure 14 specifically demonstrates the observation with the visual illustration using LBP features for character “e” and image document Lena. The texture pattern within the blank space (the area between the inner red dotted line block and the outside black solid-line block) is different from the texture pattern within the red dot block for character “e” using LBP features. However, the texture pattern using LBP features is quite consistent for the extracted block of image Lena. Under such circumstances, great effort should focus on the ROI of text document for the character itself where pattern recognition technique [37] can be applied to extract the character itself for the accuracy improvement. Similar approach should be applied for image document to avoid blank area for feature extraction. In brief, the selection of interested content from printed document should be wisely considered and processed separately for text and image to augment the influence of the region of interest for effective feature extraction.

### 3.4.4 Feature effectiveness

From Tables 9 and 10, the results show that different feature set yields different identification accuracy rates for text and image document. For example, LBP features are very suitable and promising features for image document. On the other hand, LBP filters are not significantly superior to other filters in the classification for text documents. Similarly, Gabor, GLCM, and Wiener features showed stable and effective accuracy rates for text and image documents which explains why many studies applied GLCM filters [8, 33–36, 53, 55, 56] or DWT filters [5, 53, 55, 56] as the major techniques for feature extraction. On the other hand, the performance comparison between the usage of feature selection and all features is very close for image document. Especially, LBP filter set along can almost achieve correct classification. Table 11



**Fig. 14** The result of feature extraction using LBP features for (a) character “e” and (b) image document “Lena”

summarized the accuracy rate prediction using feature fusion and selection for text and image documents. For each document we tested, four characters (text document) and also four image documents, have been classified. The average prediction results for text document is 97.38% and 99.88% for image document while all 306 features are applied. On the other hand, the accuracy prediction rate can be improved to 98.53% for text documents and 99.96% for image documents when 222 selected features are implemented. Since image patch block contains more texture information than the block from text document, it is justifiable that the results for the image document are superior to the results for the text document.

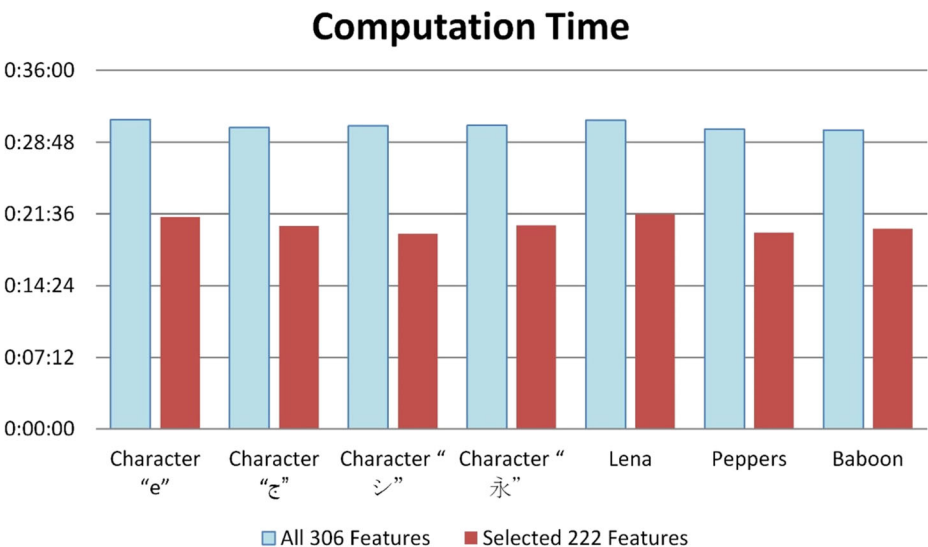
Nevertheless, it is very useful to use feature selection for better results since it is necessary to explore more features in the future to enrich the feature sets and the proposed decision-fusion model could effectively search the best features in digital forensics researches.

### 3.4.5 Computation analysis

Figure 15 illustrates the line graph for the classification computation time required in scanned image for different characters and images by feature based SVM system. The time required for both text and image documents are very close. For example, using all 306 features for classification requires around 30 min. However, using selected 222 features for classification needs around 20 min. According to Section 3.1's description, the time expense is based on ten times of training and testing in order to get the average results for classification. Therefore, in average, one cycle of training and testing only need 2 min by selected 222 features.

The experimental system is using Intel i5-2400CPU @3.10GHz, Window 7 64-bit system with 8G RAM. The whole computation software and hardware system could be further optimized to speed up the entire processes.

On the other hand, CNN based computation generally requires higher level of computation specification and graphics processing unit (GPU) is especially important. The experimental system is using Intel i7-7700CPU @3.60GHz, Window 10 64-bit system with 16G RAM.



**Fig. 15** Computation time for printer classification

GPU is GeForce GTX 1070 with 8G RAM. Table 22 tabulates the simulation results for Japanese character “あ” by different layer’s CNNs system. The training time is about 40 min for 7 layer system and 76 min for 13 layer system. As expected, the testing time is much shorter for CNNs based system than feature based SVM system. For example, the testing time is only 17.727 s for 7 layer system and 22.428 s for 13 layer system.

Due to the limit of the currently research resources, both methods are implemented in different computation systems. As mentioned above, the results in Fig. 15 and Table 22 just demonstrate the computation requirement for the classification applications. The whole system could be further improved for better hardware and software optimization. In addition, the cloud computing could be implemented to share the computation load. It may also facilitate the application for real-time identification using edge computing.

### 3.4.6 Accuracy analysis

Besides the discussion of the calculation requirement, the accuracy of the classification capability for the identification system is more critical and concerted in real applications. By averaging the results from Tables 21 and 23 respectively tabulates the comparison of the forensic classification results for scanned images by two different systems, i.e. feature based SVM system and deep learning system.

Feature based SVM system can achieve accuracy at 98.72% and CNNs can achieve accuracy at 97.7% by 7 layer system for text scanned image. Both results still exists 1 % difference. For Lena, Peppers and Baboon images, SVM based system can achieve average accuracy at 99.95% and CNNs can achieve average accuracy at 99.95% by 10 layer system. Both results are also comparable at high rates for image documents. Therefore, both systems perform quite comparable for image documents but SVM based classification system still has superior results to the outcomes from deep learning system for text documents.

In summary, even the feature based classification system generally relies on much of the efforts from dedicated feature engineering, with prior expert domain knowledge, to select the useful features. On the other hand, the advantage of deep learning reduces the human intervention and the system can extract and organize the discriminative information from the data automatically. If the initial classification decision is needed within limited time-span, it is still beneficial to use existing important features for classification in order to retrieve the best accuracy under constrained situation.

### 3.4.7 Different deep learning structure

Since the breakthrough in 2012 ImageNet competition [44], Deep Neural Network (DNN) emerges as a prominent technique for object detection and image classification at large scale. Several other DNNs with increasing complexity have been submitted to the challenge in order

**Table 22** Comparison of computation time for Japanese character “あ” by different layer’s CNNs

	1 conv 7 layers	2 conv 10 layers	3 conv 13 layers
Training time	2436 s	3915 s	4588 s
Classification rate (Testing time)	99.44% (17.727 s)	99.16% (20.315 s)	97.66% (22.428 s)

**Table 23** Comparison of average classification accuracy

Classification rate	Text Scanned image	Image Scanned image
SVM	98.72%	99.95%
CNNs	97.7% (7 layer)	99.95% (10 layer)

to achieve better performance. For example, AlexNet [25] had a very similar architecture as LeNet by LeCun et al. [26] but was deeper, with more filters per layer, and with stacked convolutional layers. The GoogLeNet [50] used a CNN inspired by LeNet but implemented a novel element which is dubbed an inception module. This module is based on several very small convolutions in order to drastically reduce the number of parameters. VGGNet [48] consists of 16 convolutional layers and is very appealing because of its very uniform architecture. Residual Neural Network (ResNet) by He et al. [13] introduced a novel architecture with “skip connections” and features heavy batch normalization.

In general, those highly acclaimed deep learning algorithms are basically designed for computer vision and image recognition with large scale labelled data samples. However, the data size of printed sources are basically small scale since the digital data need to be printed, scanned, extracted and tagged for reference, and huge human involvement is required. Therefore, it is still an issue whether the deep learning algorithms are suitable for digital forensic. Besides, extra time and works are needed to seek the best parameter settings for the deep learning structures. Since many advanced and complicated deep learning structures are generally based on CNN, the authors had implemented the model, build the structure step by step, and fine tune the parameters to get the highest accuracy for printer source identification in this study. Hence CNN model is feasible for forensic study, future researchers could applied above mentioned DNN model like GoogLeNet, VGGNet, ResNet or new structures to validate their feasibility, and improve their performance for digital forensics of printer source identification.

## 4 Conclusion

This study focused on the investigation of digital forensics for text and image document that are printed from different printer. Ten feature filter sets with 306 features are utilized by the proposed decision-theoretical model for implementation. The method emphasized on fusion of five feature selection techniques to get the most important features for the best identification accuracy rate using SVM classification of machine learning approach. Due to the influence of texture and color from printed paper, it is necessary to apply the same quality paper during the analysis in order to make fair comparison. By conducting experiments with different parameters, the proposed solution can differentiate printer sources efficiently which are all superior to the existing methods. Compared with the up-to-date technology based on deep learning, the proposed approach can achieve the best performance for input either text or image documents. For deep learning classification system, it performs comparable for image documents but the results are still inferior to the feature based SVM system for text document. While time is constrained, it is still valuable to use existing known good feature filters for initial classification and retrieve the decision with high accuracy. In summary, the proposed decision-theoretical model can be very efficiently implemented for real world digital forensic applications.



Regarding the future research, it is a continuous work to explore more useful features to expand the feature space. Feature based SVM classification and deep learning based system should be further investigated and optimized for universal applications. In addition, the efficient pattern recognition technique for text recognition can be involved for ROI partition for either text or image document. New research direction for cancer classification of medical image is creating new avenues for the development in early cancer detection. Importantly, early detection has advantages to determine the appropriate treatment to avoid the risk of errors for the patients. The techniques developed in this study could be leveraged into medical fields.

**Acknowledgments** This work was partially supported by the National Science Council in Taiwan, Republic of China, under NSC104-2410-H-009-020-MY2 and NSC106-2410-H-009-022-.

The authors would like to thank the anonymous reviewers with their valuable comments to improve the quality of this manuscript. Special thanks to Jin-Sheng Yin and Goang-Jiun Wang at National Chiao Tung University who help the revision and the software experiments.

### Appendix: Formulas of feature filters

Brief description of the formulas for ten feature filter sets is shown below:

Feature Filter	Image quality Measures	Formula
GLCM	Region of interest $R$ (ROI) GLCM	$R = \sum_{(i,j) \in ROI} 1$ $GLCM(i, j) = \frac{1}{\sum_{(i,j)} Img(i, j)} Img(i, j)$ <p>where <math>(i, j)</math> indicates the spatial location of image. <math>Img(i, j)</math> is the probability from location <math>(i, j)</math>.</p>
DWT	Dilation 3 wavelet functions	$\Psi^{(H)}(x, y), \Psi^{(V)}(x, y), \text{ and } \Psi^{(D)}(x, y),$ <p>When the wavelete function is sparable by <math>f(x, y) = f_1(x), f_2(y)</math>, then these functions rewritten to <math>\phi(x, y) = \phi(x), \phi(y)</math>  <math>\Psi^{(H)}(x, y) = \Psi(x), \phi(y)</math>  <math>\Psi^{(V)}(x, y) = \phi(x), \Psi(y)</math>  <math>\Psi^{(D)}(x, y) = \Psi(x), \Psi(y)</math>                      where <math>\Psi^{(H)}(x, y), \Psi^{(V)}(x, y)</math>, and <math>\Psi^{(D)}(x, y)</math> are called horizontal, vertical, and diagonal wavelets</p>
Gaussian	$G(x, y)$ is Gaussian matrix element at position $(x, y)$	$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$ <p>where <math>G(x, y)</math> is Gaussian matrix element at position <math>(x, y)</math>, <math>\sigma</math> is the standard deviation.</p>
LoG	$Log(x, y)$ is the high-frequency Laplacian filter	$Log(x, y) = -\frac{1}{\pi\sigma^4} \left[ 1 - \frac{1}{\pi\sigma^4} \right] e^{-\frac{1}{\pi\sigma^4}}$
Unsharp	$f_s(x, y)$ is the sharpened imaged from unsharp mask	$f_s(x, y) = f(x, y) - \bar{f}(x, y)$ <p>where <math>\bar{f}(x, y)</math> is a blurred version of <math>f(x, y)</math></p>
Wiener	$H(u, v)$ is function of Wiener filter	$g(x, y) = f(x, y) + n(x, y)$ $H(u, v) = \frac{P_f(u, v)}{P_f(u, v) + \sigma^2}$ <p>Where <math>\sigma^2</math> is variance from the noise in Eqs. (9), <math>P_f(u, v)</math> is the signal power spectrum</p>
Gabor	$sx$ & $sy$ are the variance along $x$ and $y$ axis, $f$ is the frequency of sinusoidal function and $\theta$ is the orientation of Gabor function	$G(x, y) = \frac{f^2}{\pi\eta\gamma} \exp\left(-\frac{f^2}{\pi\eta\gamma}\right) \exp(j2\pi f x' + \phi)$ $x' = x \cos(\theta) + y \sin(\theta)$ $y' = y \cos(\theta) - x \sin(\theta)$

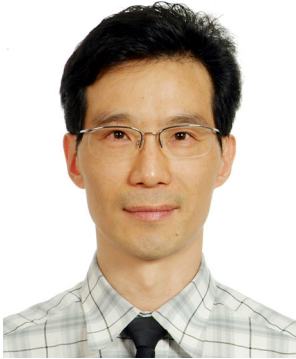
Haralick	Angular second moment	$\sum_i \sum_j p(i, j)^2$
	Contrast	$\sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \right\},  i-j  = n$
	Correlation	$\frac{\sum_i \sum_j (ij) p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$
	Sum of squares (variance)	$\sum_i \sum_j (i - \mu)^2 p(i, j)$
	Inverse different moment	$\sum_i \sum_j \frac{i}{1+(i-j)^p} p(i, j)$
	Sum average	$\sum_{i=2}^{2N_g} i p_{x+y}(i)$
	Sum variance	$\sum_{i=2}^{2N_g} (i-f_s)^2 p_{x+y}(i) - \sum_{i=2}^{2N_g} p_{x+y}(i) \log \left\{ p_{x+y}(i) \right\}$
	Sum entropy	$-\sum_i \sum_j p(i, j) \log(p(i, j))$
	Entropy	variance of $p_{x-y}(i)$
	Difference variance	$-\sum_{i=0}^{N_g-1} p_{x-y}(i) \log \left\{ p_{x-y}(i) \right\}$
	Difference entropy	$\frac{HXY - HXY_1}{\max\{HXY, HXY_1\}}$
	Info. measure of correlation 1	$(1 - \exp[-2.0(HXY2 - HXY)])^{\frac{1}{2}}$
	Info. measure of correlation 2	$HXY1 = -\sum_i \sum_j p(i, j) \log(p(i, j))$
	Max. correlation coefficient	where $HX$ and $HY$ are the entropies of $p_x$ and $p_y$ , $HXY1 = -\sum_i \sum_j p(i, j) \log \{ p_x(i) p_y(j) \}$ , and $HXY2 = -\sum_i \sum_j p_x(i) p_y(j) \log \{ p_x(i) p_y(j) \}$ (the second largest eigenvalue of $Q$ ) <sup>1/2</sup> where $Q(i, j) = \sum_k \frac{p(i, k) p(j, k)}{p_x(i) p_y(k)}$
Fractal	$\Delta(x, y)$ : fractal feature vector	$\Delta(x, y) = \{ 1, 0, \text{if } \exists (x', y') \in N_4[x, y], Ib(x', y') = 0 \}$ $Ib(x', y') = 1, \text{ otherwise}$ where $N_4[x, y]$ is the set of pixels that are 4-connected to $(x, y)$ from the image. $\Delta(x, y)$ uses the value 1 if the pixel at position $(x, y)$ in the binary image $Ib(x, y)$ that has the value 1 and having one neighboring pixel with the value 0. Otherwise, $\Delta(x, y)$ takes the value 0.
LBP	$LBP_{P,R}(x_c, y_c)$ LBP features where $P$ sampling points on a circle of $R$ radius	$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p s(x) = \{ 1, \text{if } x \geq 0; 0, \text{otherwise.}$

References

1. Ali GN, Mikkilineni AK, Chiang PJ, Allebach GT, Delp EJ (2003) Intrinsic and extrinsic signatures for information hiding and secure printing with electrophotographic devices. In International Conference on Digital Printing Technologies. New Orleans, LA, USA; 28 Sept–3 Oct, 511–515
2. Bekhti MA, Kobayashi Y (2016) Prediction of vibrations as a measure of terrain traversability in outdoor structured and natural environments. In: Image and video technology, Vol. 9431 of the series lecture notes in computer science. Springer International Publishing, Auckland 282–294. [https://doi.org/10.1007/978-3-319-29451-3\\_23](https://doi.org/10.1007/978-3-319-29451-3_23)
3. Bulan O, Mao J, Sharma G (2009) Geometric distortion signatures for printer identification. International conference on acoustics, speech and signal processing (ICASSP), Taipei, 1401–1404. <https://doi.org/10.1109/ICASSP.2009.4959855>
4. Burger W, Burge MJ (2018) Digital image processing: an introduction algorithmic using Java. Springer Science Business Media, New York
5. Choi JH, Lee HY, Lee HK (2013) Color laser printer forensic based on noisy feature and support vector machine classifier. Multimed Tools Appl 67:363–382. <https://doi.org/10.1007/s11042-011-0835-9>
6. Costa AF, Humpire-Mamani G, Traina AJM (2012) An efficient algorithm for fractal analysis of textures. SIBGRAPI conference on graphics, patterns and images, August, Ouro Preto, 39–46. <https://doi.org/10.1109/SIBGRAPL.2012.15>
7. Daugman JG (1988) Complete discrete 2D Gabor transforms by neural networks for image-analysis and compression. IEEE Trans Acoust Speech Signal Process 36(7):1169–1179. <https://doi.org/10.1109/29.1644>

8. Ferreira A, Navarro LC, Pinheiro G, Santos JAD, Rocha A (2015) Laser printer attribution: exploring new features and beyond. *Forensic Sci Int* 247:105–125. <https://doi.org/10.1016/j.forsciint.2014.11.030>
9. Gonzales RC, Woods RE (2008) *Digital image processing*, 3rd edn. Prentice Hall, New Jersey
10. Gonzales RC, Woods RE, Eddins SL (2009) *Digital image processing using MATLAB*, 2nd edn. Gatesmark, United States
11. Haghighat M, Zonout S, Abdel-Mottaleb M (2015) CloudID: trustworthy cloud-based and cross-enterprise biometric identification. *Expert Syst Appl* 42(21):7905–7916. <https://doi.org/10.1016/j.eswa.2015.06.025>
12. Haralick RM, Shanmugam K, Dinstein I (1973) Textural features for image classification. *IEEE Trans Syst Man Cybernet SMC* 3(6):610–621
13. He K et al (2016) Deep residual learning for image recognition. *IEEE conference on computer vision and pattern recognition (CVPR)* 770–778
14. Hinton G, Salakhutdinov R (2006) Reducing the dimensionality of data with neural networks. *Science* 313(5786):504–507
15. Hsu CW, Chang CC, Lin CJ (2003) A practical guide to support vector classification. National Taiwan University, Taipei <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>. Accessed 8 Apr 2017
16. <http://www.explainthatstuff.com/laserprinters.html>. Accessed 2 Apr 2017
17. <http://computer.howstuffworks.com/laser-printer2.htm>. Accessed 2 Apr 2017
18. Hubel D, Wiesel T (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol* 160(1):106–154
19. Jurić I, Randelović D, Karlović I, Tomić I (2014) Influence of the surface roughness of coated and uncoated papers on the digital print mottle. *J Graph Eng Des* 5(1):17–23
20. Kawasaki M, Ishisaki M (2009) Investigation into the cause of print mottle in halftone dots of coated paper: effect of optical dot gain non-uniformity 63(11):1362–1373. <http://www.tappi.org/content/061PGA/5-4%20Kawasaki%20M%20Ishisaki.pdf>. Accessed 7 April 2017
21. Kee E, Farid H (2008) Printer profiling for forensics and ballistics. *ACM Workshop on Multimedia and Security*, 3–10
22. Khanna N, Delp EJ (2010) Intrinsic signatures for scanned documents forensics: effect of font shape and size. *Proceedings of 2010 I.E. international symposium on circuits and systems (ISCAS)*, 30 May– 2 June. <https://doi.org/10.1109/ISCAS.2010.5537996>
23. Kim DG, Lee HK (2014) Color laser printer identification using photographed halftone images. *Proc. of EUSIPCO*. September, IEEE, Lisbon, 795–799
24. Kim KI, Jung K, Park SH, Kim HJ (2002) Support vector machines for texture classification. *IEEE Trans Pattern Anal Mach Intell* 24(11):1542–1550. <https://doi.org/10.1109/TPAMI.2002.1046177>
25. Krizhevsky A, Sutskever I, Hinton G (2012) ImageNet classification with deep convolutional neural networks. *Process Int Conf Neural Inf Process Syst (NIPS)* 1:1097–1105
26. Lecun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. *Proc IEEE* 86(11):2278–2324
27. Lewis JA (2014) *Forensic document examination: fundamentals and current trends*. Elsevier, Oxford. <https://doi.org/10.1016/B978-0-12-416693-6.12001-6>
28. Lin CJ (2007) A tutorial of the wavelet transforms. National Taiwan University, <http://disp.ee.ntu.edu.tw/tutorial/WaveletTutorial.pdf>. Accessed 3 Apr 2017
29. Lopez FM, Martins DC, Cesar RM (2008) Feature selection environment for genomic applications. *BMC Bioinformatics* 9:451. <https://doi.org/10.1186/1471-2105-9-451>
30. Mäenpää T, Pietikäinen M (2004) Texture analysis with local binary patterns. In: Chen CH, Wang PSP (eds) *Handbook of pattern recognition & computer vision*, 3rd edn. World Scientific, Singapore, pp 115–118
31. Markoff J (2012) How many computers to identify a cat? 16,000. *The New York*. Retrieved June 22, 2012, from <https://mobile.nytimes.com/2012/06/26/technology/in-a-big-network-of-computers-evidence-of-machine-learning.html>
32. McAndrew A (2016) *A computational introduction to digital image processing*. CRC Press, Boca Raton
33. Mikkilineni AK, Chiang PJ, Ali GN, Chiu GT, Allebach JP, Delp EJ (2004) Printer identification based on textural features. *Intl. conference on digital printing technologies* 306–311
34. Mikkilineni AK, Chiang JP, Ali GN, Chiu GT, Allebach JP, Delp EJ (2005) Printer identification based on graylevel co-occurrence features for security and forensic applications. *Intl. Conference on Security, Steganography and Watermarking of Multimedia Contents VII*, Proc. SPIE. 5681, 430–440, March 21. <https://doi.org/10.1117/12.593796>
35. Mikkilineni AK, Arslan O, Chiang PJ, Kumontoy RM, Allebach JP, Chiu GT (2005) Printer forensics using SVM techniques. *Intl. conference on digital printing technologies*, 223–226
36. Mikkilineni AK, Khanna N, Delp EJ (2010) Texture based attacks on intrinsic signature based printer identification. *Proceedings SPIE 7541, Media Forensics and Security II*, 28 January. <https://doi.org/10.1117/12.845377>

37. Netzer Y, Wang T, Coates A, Bissacco A, Wu B, Ng A (2011) Reading digits in natural images with unsupervised feature learning. NIPS workshop on deep learning and unsupervised feature learning
38. Ojala T, Pietikäinen M, Harwood D (1996) A comparative study of texture measures with classification based on featured distributions. *Pattern Recogn* 29(1):51–59. [https://doi.org/10.1016/0031-3203\(95\)00067-4](https://doi.org/10.1016/0031-3203(95)00067-4)
39. Ojala T, Pietikäinen M, Mäenpää T (2002) Multiresolution gray-scale and rotation invariant texture classification with LBP. *IEEE Trans Pattern Anal Mach Intell* 24(7):971–987. <https://doi.org/10.1109/TPAMI.2002.1017623>
40. Pudil P, Ferry FJ, Novovicova J, Kittler J (1994) Floating search methods for feature selection with nonmonotonic criterion functions. *IEEE*, 1051-465U9, <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=576920>. Accessed 3 Apr 2017
41. Pudil P, Novovicova J, Kittler J (1994) Floating search methods in feature selection. *Pattern Recogn Lett* 15: 1119–1125
42. Qiu Z, Jin J, Lam HK, Zhang Y, Wang X, Cichocki A (2016) Improved SFFS method for channel selection in motor imagery based BCI. *Neurocomputing*. <https://doi.org/10.1016/j.neucom.2016.05.035>
43. Rumelhart E, Geoffrey E, Ronald J (1986) Learning representations by back-propagating errors. *Nature* 323:533–536
44. Russakovsky O et al (2015) Imagenet large scale visual recognition challenge. *Int J Comput Vis* 115(3): 211–252
45. Ryu SJ, Lee KY, Im DH, Choi JH, Lee HK (2010) Electrophotographic printer identification by halftone texture analysis. In: *IEEE Intl. conference on acoustics speech and signal processing (ICASSP)*, 1846–1849. <https://doi.org/10.1109/ICASSP.2010.5495377>
46. Say OT, Sauli Z, Retnasamy V (2013) High density printing paper quality investigation. *IEEE Regional Symposium on Micro and Nano electronics (RSM)*, Langkawi, 273–277. <https://doi.org/10.1109/RSM.2013.6706528>
47. Schalkoff RJ (1989) *Digital image processing and computer vision*. Wiley, Australia
48. Simonyan K. and Zisserman A. (2015) Very deep convolutional networks for large-scale image recognition. *IEEE conference on computer vision and pattern recognition (CVPR)*, arXiv preprint arXiv:1409.1556
49. Su R, Pekarovicova A, Fleming PD, Bliznyuk V (2005) Physical properties of LWC papers and Gravure Ink Mileage. [https://www.researchgate.net/publication/251423637\\_Physical\\_Properties\\_of\\_LWC\\_Papers\\_and\\_Gravure\\_Ink\\_Mileage](https://www.researchgate.net/publication/251423637_Physical_Properties_of_LWC_Papers_and_Gravure_Ink_Mileage). Accessed 3 Apr 2017
50. Szegedy, C. (2015) Going deeper with convolutions. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. <https://doi.org/10.1109/CVPR.2015.7298594>
51. The Electron Microscope (2017) <http://www.microscopemaster.com/electron-microscope.html>. Accessed 11 Apr 2017
52. Tong S, Koller D (2001) Support vector machine active learning with applications to text classification. *J Mach Learn Res* 45–66. <http://www.jmlr.org/papers/volume2/tong01a/tong01a.pdf>. Accessed 7 Apr 2017
53. Tsai MJ, Liu J (2013) Digital forensics for printed source identification. In: *IEEE international symposium on circuits and systems (ISCAS)*, May, 2347–2350. <https://doi.org/10.1109/ISCAS.2013.6572349>
54. Tsai MJ, Liu J, Wang CS, Chuang CH (2011) Source color laser printer identification using discrete wavelet transform and feature selection algorithms. *IEEE international symposium on circuits and systems (ISCAS)*, May, Rio de Janeiro, 2633–2636. <https://doi.org/10.1109/ISCAS.2011.5938145>
55. Tsai MJ, Yin JS, Yuadi I, Liu J (2014) Digital forensics of printed source identification for Chinese characters. *Multimed Tools Appl* 73:2129–2155. <https://doi.org/10.1007/s11042-013-1642-2>
56. Tsai MJ, Hsu CL, Yin JS, Yuadi I (2015) Japanese character based printed source identification. *IEEE International Symposium on Circuits and Systems (ISCAS)*, May, Lisbon, 2800–2803. <https://doi.org/10.1109/ISCAS.2015.7169268>
57. Vega LR, Rey H (2013) *A rapid introduction to adaptive filtering*. Springer-Verlag, Berlin
58. Wu Y, Kong X, You X, Guo Y (2009) Printer forensics based on page document’s geometric distortion. *Intl. conference on image processing (ICIP)*, Cairo, 2909–2912. <https://doi.org/10.1109/ICIP.2009.5413420>
59. Zhou H, Wu J, Zhang J (2010) *Digital image processing: part 1*. Ventus Publishing ApS, Denmark



**Min-Jen Tsai** received the B.S. degree in electrical engineering from National Taiwan University in 1987, the M.S. degree in industrial engineering and operations research from University of California at Berkeley in 1991, the engineer and Ph.D. degrees in Electrical Engineering from University of California at Los Angeles in 1993 and 1996, respectively. He served as a second lieutenant in Taiwan army from 1987 to 1989. From 1996 to 1997, he was a senior researcher at America Online Inc. In 1997, he joined the institute of information management at the National Chiao Tung University in Taiwan and is currently a full professor. His research interests include multimedia system and applications, digital right management, digital watermarking and authentication, digital forensic, enterprise computing for electronic commerce applications. Dr. Tsai is a member of IEEE, ACM and Eta Kappa Nu.



**Imam Yuadi** received his Ph.D. degree from the Institute of Information, National Chiao Tung University in 2017. He is currently an assistant professor in the Department of Information and Library Science, Airlangga University, Indonesia. Yuadi studied in the Library and Information Science from Padjadjaran University in 1999 and received his master degree in Information Technology Management from Sepuluh Nopember Institute of Technology at Surabaya in 2009. His research interest includes digital forensics, image recognition, computer vision and digital library.



**Yu-Han Tao** is currently a Ph.D. student at the institute of information management, National Chiao Tung University.