2018 24th International Conference on Pattern Recognition (ICPR)
Beijing, China, August 20-24, 2018

Document Images Watermarking for Security Issue using Fully Convolutional Networks

Cu Vinh Loc, Jean-Christophe Burie and Jean-Marc Ogier
L3i Laboratory, La Rochelle University, France
E-mail: {vinh_loc.cu, jean-christophe.burie, jean-marc.ogier}@univ-lr.fr

***Abstract*—In the literature, the watermarking schemes for document images in spatial domain mainly focus on text content, so they need to be further improved to be possibly applied on general content. In this paper, we propose a blindly invisible watermarking approach for security matter of general grayscale documents. In order to detect stable regions used for hiding secret data, we make the best use of fully convolutional networks (FCN). The FCN for the problem of document structure segmentation is adjusted to solve the problem of watermarking regions detection wherein we consider various types of segmented content regions having the same label. The segmented content regions are then known as watermarking regions. Next, the watermarking pattern is constructed with the aim of detecting potential positions where the watermarking process is carried out. Lastly, the watermarking algorithm is developed by dividing gray level values pertaining to each watermarking pattern into two groups for carrying one watermark bit. The experiments are performed on various document contents, and our approach obtains high performance in terms of imperceptibility, 不可感知性 capacity and robustness against distortions caused by JPEG compression, geometric transformation and print-and-scan process.**

I. Introduction

Due to the convenience and popularity of digital information exchange daily, the legal documents are mostly scanned and stored under digital format as document images. These documents are in use at governmental agencies, banks, educational institutions, military sites, etc. where the protection of document security against unauthorized access is not just a matter that draws much attention of those who have proper permission. The securing of documents also poses many challenges which are being tackled by researchers, especially in the field of document analysis and recognition. To tackle these problems, digital watermarking can be used as an interesting solution that enables to hide a secret message or signal into the digital medium. The digital medium mentioned in this paper is a document image or a host document. The secret message or signal is called watermark. After hiding a watermark, the host document becomes a watermarked document. Moreover, the watermark is hidden into a host document such that the distortions caused by the hiding process are less to be perceptible, and it is also robust enough to withstand common distortions.

Most existing watermarking approaches have been proposed for text content [1]–[7] except recent works [8], [9] designed for general content but they have been developed in the transform domain. The methods for text content seem not to work well for general content because on their work no detail is given on how to identify and separate text and non-text elements. In addition, the approaches for natural images [10]–[14] can be applied for document images but they need to have further adjustments to improve the robustness (e.g. such extracted features as keypoints or edges are unstable with regard to distorted document images) or to eliminate empty regions (e.g. document images often contain a lot of empty areas). Besides, at the time of performing this work, we have found some watermarking approaches for natural images using convolutional neural network (CNN) [15], [16] but have not found FCN-based watermarking approaches.

Despite our previous method [17] has improved the performance compared to the existing approaches, the extracted regions for watermarking process is still unstable against high noises. Thus, we propose a FCN-based approach for general document images and introduce this advanced technique to the field of document watermarking for security concern. Different from the CNN-based approaches wherein the authors leverage weight parameters of learning framework for watermarking process on the fixed size images, here we train the FCN so that the trained network can be used for generating a salient map describing watermarking regions of an arbitrary sized document. The process of our watermarking operations is separately designed with the phase of training network.

The rest of the paper is organized as follows. We present related works in the next section. In section III, the proposed method is detailed. Experimental results are presented in section IV. Section V concludes our works.

II. Related work

A. *Watermarking methods for document images*

Palit and Garain [1] proposed a method by hiding data into prototypes constructed from document. The authors use pattern matching techniques for clustering symbols into prototypes, and the image of prototypes is represented under binarized form. This method can withstand distortions like JPEG compression, uniform random noise and scaling. However, it is designed for Indian text document. Kim and Oh [2] proposed an approach based on edge direction histograms. Sobel edge operator is used to compute the edge direction histograms. The document is partitioned into blocks where the histogram of text block is adjusted for carrying watermark bits. This approach is robust to such distortions as blurring, sharping, rotation, etc. Hu et al. [4] proposed a method in which the document is divided into partitions by using support vector machine.

Data is then hidden into uniform partitions by making some small modifications such that some lines in partition are suitable for hiding watermark bits. Kim et al. [3] put forward word classification and inter-word space statistics for text document. The watermark bits are hidden by slightly shifting words to left or right in each segment defined by grouping consecutive words in a line. This method gives good quality and robustness. Zou and Shi [5] proposed a method depending on inter-word spaces. The inter-word spaces between adjacent words in a text row are detected and divided into two sets. Each set has equal number of inter-word space, data hiding is performed by altering these spaces. This approach is robust to print-photocopy-scan (PCS) noise.

Lina et al. [7] proposed a method based on stroke direction. Data bits are hidden into individual characters by adjusting the directional characteristic of rotatable strokes. This approach gives high capacity, imperceptibility and resistance to PCS noise. However, it can only be applied for Chinese text. Varna et al. [6] proposed an approach wherein the data is hidden into the vertically left edge of the character stroke by adding or deleting two groups of pixels. The authors combined message packetization and error correction code for enhancing the performance of their approach. This method is robust to PCS noise. However, the distortion caused by hiding process is likely perceptible. More recently, Chetan and Nirmala [9] proposed integer wavelets and block coding of binary watermarks. This method is likely resistant to common noises. Another work for hiding watermark into frequency domain is put forward by Horng et al. [8]. This method is based on discrete cosine transform (DCT), singular value decomposition (SVD) and Genetic algorithm (GA). The authors show that their method is robust to several types of distortion. Most recently, we proposed another method [17] based on stable regions and object filling for general document. The stable regions are detected by combining image processing operations and non-subsampled contourlet transform. The object filling is used as referential information for watermarking process. This method is robust to JPEG compression, geometric transformation and print-and-scan noise.

B. *Watermarking methods for natural images*

Amiri et al. [11] proposed a method based on discrete wavelet transform (DWT)-DCT. One-dimensional DCT is used to extract the final coefficients. Two-dimensional DWT is applied to the original image to obtain the mid-frequency subbands. Besides, the authors also apply GA and modeling algorithm for their work. This scheme is robust to print-and-scan distortion. Alia et al. [10] proposed another scheme based on DWT and DCT. The authors used two DCT-transformed sub-vectors for hiding watermark bits. The triangular regions and Zernike moments-based approach is proposed by Munib et al. [13] in which Harris detector is used to extract keypoints. Delaunay tessellation is then employed to divide image into distinct triangular segments on these extracted feature points. The watermark is hidden into the magnitude of Zernike moments of each triangle segment. Zolotavkin et al. [12] proposed a technique of Distortion Compensation (DC) for two dimensional Quincunx Lattice Quantization. The parameter for controlling DC is modified to improve the efficiency of their scheme. An adaptive blind watermarking in DWT and DCT is proposed by Maedeh et al. [14]. The authors use fuzzy system to compute strength factors which are relied upon image features like edge concentration, saliency and intensity. More recently, Mun et al. [16] put forward a scheme based on CNN. The authors divide the host image, watermark into non-overlapping blocks and take advantage of CNN's weight parameters during training phase for watermarking process. Each block can be capable of carrying one watermark bit.

III. Proposed method

A. *FCN-based detection of watermarking regions*

In reality, the document images are represented under several forms with various size. After watermarking process, the watermarked document is expected to keep a similar shape with the host document. The FCN [18] is adopted for developing our watermarking scheme because this network is able to take an arbitrary sized document and generate an output with the same dimension (since it has no fully connected layer). The FCN is initially designed for semantic image segmentation, and this network is very powerful to perform the document structure segmentation. Our observations have shown that the segmented content regions of document such as running text, headline, picture, table, etc. are stable against distortions. Thus, we transform the problem of document structure segmentation including many labels describing various segmented content regions to the problem of watermarking regions detection including two labels: one describes background region (black color in Fig. 2(b)), and another depicts all segmented content regions (white color in Fig. 2(b)). It means that the segmented content regions have the same label (e.g. a segmented region containing the running text does not differentiate from other segmented content regions and so on).

Our network (Fig. 1) is based on [18] in which we use the VGG 16-layer [19] network and replace three fully connected layers with convolutional layers (two blocks of additional layers) for preserving spatial information. The feature maps generated in the phase of convolution operations (downsampling) are reduced in dimension, so they need to be reconstructed by performing upsampling through transpose convolution. This task produces the feature maps with the same dimension of original document in the output layer. Besides, we apply softmax layer on the output to transform the result of network into a two-class problem for representing the probability of document's watermarking regions. Next, we describe briefly

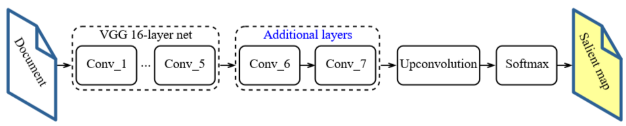


Fig. 1. The architecture of FCN for detecting watermarking regions.

our network as follows. The first five convolutional blocks are analogous to the VGG-16 net in which the first two blocks (Conv_1, Conv_2) contain two convolutional layers in each, while the next three blocks (Conv_3, Conv_4, Conv_5) contain three convolutional layers in each. The kernel size used in the first five convolutional blocks is set to 3×3. Each additional block (Conv_6, Conv_7) contains two convolutional layers including one with a kernel size of 3×3 and another with a kernel size of 2×2. The step size used for shifting convolution kernel is of 1×1. The activation function used in all layers of network is a rectified linear unit (ReLU). Besides, the max pooling layer with a kernel size and stride of 2×2 is applied after the last convolutional layer of each block except block of Conv_7. We apply dropout layer after Conv_6 and Conv_7 for preventing the problem of overfitting. The phase of reconstructing the spatial dimension of original document is represented by "Upconvolution" block (upsampling operations) in Fig. 1 wherein transpose convolution is used for retaining the spatial information.

The annotating documents are performed by creating ground truth annotation of document's content regions. The input documents along with their ground truth described in Fig. 2 are used to feed the network for training purpose. From observation, we have seen that the feature maps generated from low blocks of convolutional layers represent overall shape of document content while regions-specific information at high blocks. The salient map is computed based on the scores of output feature maps obtained from our trained network, and it is represented under grayscale form as in Fig. 3(a). Based on this salient map, the bounding boxes (blue rectangles in Fig. 3(b)) surrounding content regions of document are easily detected by utilizing connected components. These regions are known as watermarking regions.

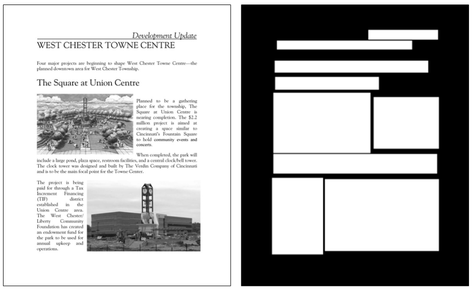


Fig. 2. The illustration of document annotation with two labels: (a) A form of general document. (b) The annotating document remarks content regions.

B. Watermarking process

1) Standardization of the input document: The watermarked document is possibly subjected to distortions caused by the printing and scanning process. The distortions could make the watermarked document rotating a certain angle or changing its dimension due to scanning with different resolutions. Thus, these distortions need to be minimized by transforming the document into a standard form prior to conducting watermarking operations. We adopt BRISK [20] detector for keypoints detection. These keypoints are used for determining the minimum rectangle surrounding the content of entire document. The rotation angle θ is then estimated relied on this minimum rectangle.

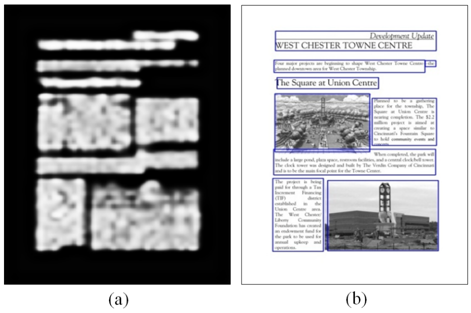


Fig. 3. Salient map and watermarking regions: (a) The salient map describes content regions of document. (b) Blue rectangles are bounding boxes.

Next, the scale factor s is calculated depending on the Euclidean distance d_e of such two points as a top left point P_1 and an intersecting point P_2 of two diagonal lines of this rectangle.

$$s = \frac{c}{d_e} \tag{1}$$

Finally, the input document image I is transformed into the standardized form I' by:

$$(x', y') = T(x_c, y_c) + sR(\theta)(x, y) \tag{2}$$

where (x_c, y_c) is the center point of standard document, (x', y') is a point of standard document I' corresponding to a point (x, y) of original document I. R and T represent the rotation and translation operation respectively.

2) Watermark hiding: The general hiding process is illustrated in Fig. 4. The watermark hiding process basically comprises of the following main steps:

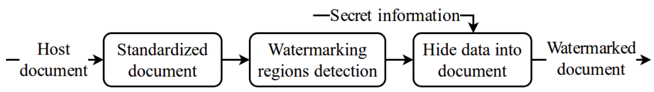


Fig. 4. The general steps of watermark hiding process.

a) Step 1 - Transform the input document into standardized form: As presented in Section III-B1

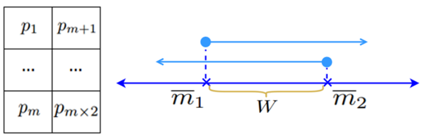
b) Step 2 - Watermarking regions detection: The watermarking regions are obtained by Section III-A. The possibilities of extracted watermarking regions consist of three cases such as: (c_1) The nested regions are eliminated because if we hide watermark bits into a nested region, the hidden data in inner region will be overwritten by the data hiding operations carried out in an outer region. Thus, only the outermost region is kept; (c_2) The small regions that are not stable against noises and need to be removed; (c_3) The overlapping regions also overwrite part of previously hidden data, so the intersecting part of a smaller region is combined into a larger one.

c) Step 3 - Identify watermarking positions: Based on the natural shape of letters, we use a watermarking pattern B with a size of $m \times 2$ (Fig. 5(a)) to scan through each watermarking region for locating suitable positions where the watermarking process will be conducted. As we know that there are many positions with high grey values interlaced either inside individual letters or among them. These positions are unusable for watermarking process, so they must be eliminated. To make sure that the watermarking process is not fallen into these positions, all gray level values $(p_1, p_2, .., p_{m\times2})$ of the watermarking pattern have to be less than a threshold δ.

d) Step 4 - Hiding watermark bits into document: We divide the gray level values of watermarking pattern B_i into two groups of m values such as $(p_1, ..., p_m)$ and $(p_{m+1}, ..., p_{m\times2})$. The idea of hiding a message bit into document is based on mean values $(\overline{m}_1, \overline{m}_2)$ corresponding to two groups of pixel values of B_i and the absolute difference d_i between these mean values. It is possibly either $\overline{m}_1 \leq \overline{m}_2$ or $\overline{m}_1 > \overline{m}_2$, and we use $\overline{m}_1 \leq \overline{m}_2$ for carrying bit 0 and $\overline{m}_1 > \overline{m}_2$ for bit 1.

$$\begin{cases} \overline{m}_1 = \frac{1}{m}\sum_{k=1}^{m} p_k; \overline{m}_2 = \frac{1}{m}\sum_{k=m+1}^{m\times2} p_k \\ d_i = |\overline{m}_1 - \overline{m}_2| \end{cases} \tag{3}$$

However, when the watermarked documents go through noises, the mean value $\overline{m}_1$ might move passed towards the mean value $\overline{m}_2$ and vice versa as illustrated in Fig. 5(b). Thus, these two means need to be separated by a certain width of W for improving the accuracy of watermark detection. A watermarking pattern B_i carries a watermark bit wm_i by:



(a) (b)

Fig. 5. (a) The watermarking pattern B_i with size of $m \times 2$. (b) The distribution of mean values $\overline{m}_1$ and $\overline{m}_2$ calculated from B_i.

- if watermark bit $wm_i = 0$:
  - $\overline{m}_1 < \overline{m}_2$ and $d_i \geq W$: the gray level values in B_i remain unchanged.
  - Otherwise: adjust the gray level values in B_i such that $\overline{m}_1 < \overline{m}_2$ and $d_i \geq W$.
- if watermark bit $wm_i = 1$:
  - $\overline{m}_1 > \overline{m}_2$ and $d_i \geq W$: the gray level values in B_i remain unchanged.
  - Otherwise: adjust the gray level values in B_i such that $\overline{m}_1 > \overline{m}_2$ and $d_i \geq W$.

The adjustment is done by: (i) permuting the first group of gray level values $(p_1, ..., p_m)$ with the second groups of gray level values $(p_{m+1}, ..., p_{m\times2})$; or (ii) modifying gray level values of the first or second group; or both of (i) and (ii).

3) Watermark detection: The watermark detection is carried out in a similar way as the watermark hiding process, just different in extracting hidden data. The watermark bit wm_i is extracted by:

$$wm_i = \begin{cases} 0, & \text{if } \overline{m}_1 \leq \overline{m}_2 \\ 1, & otherwise \end{cases} \tag{4}$$

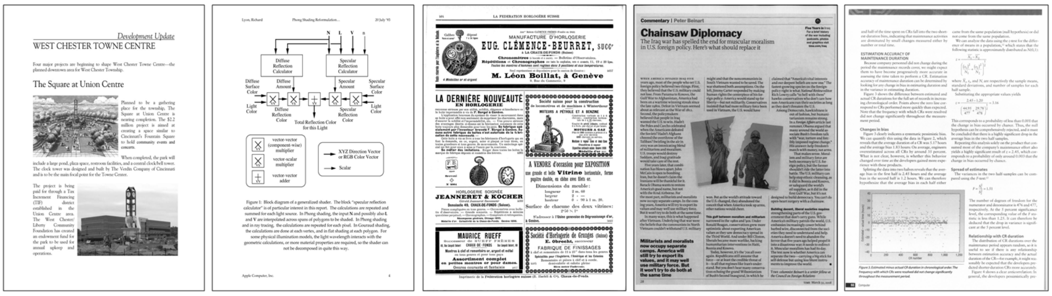
IV. Experimental Results

For training configurations, we use two datasets such as: PRImA [21] for ICDAR page segmentation competitions which is based on comprehensive and detailed representation of both simple and complex layouts. It consists of 54 sample pages of magazines together with ground truth metadata; DSSE-200 [22] contains 200 pages from magazines and academic papers. This dataset provides both appearance-based and semantics-based labels. In the context of our work, we expect rotation invariance, scale variation and variation of quality factor of compression. Thus, we generate 5080 from totally 254 document images for training samples. Regarding initialization of network parameters, the number of learning steps is set to 200000. The high momentum is assigned to 0.9. The weight decay is 5×10^{-4}. The learning rate is set to 10^{-4}, and it is adjusted to 10^{-5} when reaching half of the learning steps. The dropout rate is assigned to 0.5.

For watermarking configurations, we select 15 documents from DSSE-200 (Type-1). 60 documents are selected from L3iDocCopies [23] including: 15 documents scanned from Konica Minolta Bizhub 223 at the resolution of 300 dpi (Type-2) and 15 documents with 600 dpi (Type-3); 15 documents scanned from Fujitsu fi 6800 at the resolution of 300 dpi (Type-4) and 15 documents with 600 dpi (Type-5). We have totally tested our approach on 75 various documents. The watermark used in this experiment is "watermarking-information", this text message is modulated into 192 bits. The distance for separating the mean values is set with $W = 25$. The size of watermarking pattern is assigned to $m = 3$. These values are experimentally chosen to provide a good tradeoff among robustness, imperceptibility and capacity (if W and m are too small, the robustness will be diminished. If W is large, the imperceptibility will be degraded. If m is too large, the capacity will be low). With the sample documents in Figure 6, the value of c is set to 1836, 1765, 2414, 2147 and 4223 for (a)-(e) respectively, based on the size of minimum rectangle of entire document. The performance of our scheme is evaluated depending on the following factors.

A. Imperceptibility and capacity

The quality of watermarked document is evaluated by measuring the difference between original and watermarked document. Peak-Signal-to-Noise Ratio (PSNR) has been used in this work. To measure the imperceptibility, the watermark bits used in this part is randomly generated depending on the maximum number of bits that the host documents can contain. The PSNR, capacity of sample documents and the average values of 75 documents are demonstrated in Table I. Figure 7 shows an example of hiding 132 random message bits into host document with size of 208×186, the PSNR of this watermarked document is 46.63. We can see hidden positions (white vertical lines in Figure 7(c)) where the pixel



(a)-Doc1 (b)-Doc2 (c)-Doc3 (d)-Doc4 (e)-Doc5

Fig. 6. Sample general documents with various content: (a), (b) and (c) with corresponding size of 2550×3300, 2550×3300 and 2808×3942 represent DSSE-200 dataset. (d) and (e) with size of 2487×3513 and 4960×7015 represent L3iDocCopies dataset at the resolution of 300 and 600 dpi

(a) (b) (c)

Fig. 7. (a)-(c) are host, watermarked document and adjusted positions

values are altered. For evaluation of our scheme's performance in the environment without distortions, we hide the mentioned watermark into these 75 documents. As a result, there is no watermarked document get failed in the watermark detection.

TABLE I
IMPERCEPTIBILITY AND CAPACITY

Documents	PSNR (original vs watermarked)	Capacity (bits)
1 (Doc1)	44.60	16,115
2 (Doc2)	46.59	3,417
3 (Doc3)	44.60	15,711
4 (Doc4)	42.58	16,962
5 (Doc5)	44.26	9,533
...	...	...
Average	**43.02**	**9,167**

B. *Robustness evaluation*

1) Robustness against JPEG compression and geometric transformation: The results of watermark detection are shown in Table II in which the watermarked documents are suffered from popular distortions. We present the robustness of our scheme in detail for five sample documents. The bit error rate (BER) in Table II refers to the number of detected error bits divided by the total number of hidden bits, and we can see that the accurate ratio is degraded regarding documents scanned at low resolution. The inaccuracy of watermark detection is mainly caused by pixel values change due to noise, and this change leads to lose the integrity of mean values $\overline{m}_1$ and $\overline{m}_2$. Meanwhile, the positions of bounding boxes corresponding to the document's watermarking regions have little change.

2) Robustness against print-and-scan (PS) distortion: To evaluate our approach on this distortion, we use machine Kyocera TASKalfa 3252ci for printing the watermarked documents at the default resolution of 600 dpi. The same machine is again used for scanning the printed documents. The scanning resolution is sequentially selected to be equal to the resolutions of 600, 400, 300 and 200 dpi. As a result, when printing and scanning watermarked documents with high resolution of 600 dpi, our approach gives good result. In case of scanning the watermarked document at lower resolution, the BER is considerably increased because of the degradation of document quality. From the experiment, we have seen that the scanned documents are always subjected to geometric distortions. Figure 8 shows the average BER of five types of watermarked document at corresponding resolution.

TABLE II
THE ACCURATE RATIO OF WATERMARK DETECTION

Distortions	Bit Error Rate				
	Doc1	Doc2	Doc3	Doc4	Doc5
Rotation 5° (a)	0	0	0	0	0
Rotation 7°	0.16	0.18	0.13	0.21	0.17
Scaling 0.8	0	0	0	0.28	0
Scaling 0.9 (b)	0	0	0	0	0
Scaling 1.1 (c)	0	0	0	0	0
Scaling 1.2 (d)	0	0	0	0.25	0
(a) + (b)	0	0	0	0	0
(a) + (c)	0	0	0	0	0
(a) + (d)	0.27	0.23	0.29	0.31	0.26
JPEG 60%	0	0	0	0	0
JPEG 50%	0	0	0	0.24	0
JPEG 40%	0.17	0.16	0.21	0.34	0.25

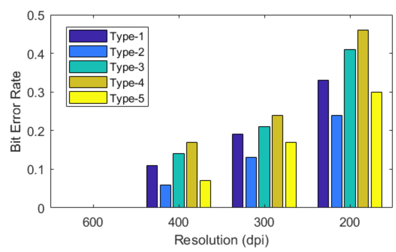


Fig. 8. The average results of watermark detection on print-and-scan noise

C. Comparison with other techniques

We define various types of distortion for our comparison including: (1) geometric transformation; (2) JPEG compression; (3) PS noise; (4) PCS noise; (5) cropping, salt and pepper noise; (6) filtering noises. Unfortunately, we have not found the dataset as well as the implementation of typical works that we compare with. Thus, we can not make a quantitative comparison of these methods on our sample documents, or make a quantitative comparison of our approach on their dataset. As a result, our comparison is performed relying on scheme's functionalities. The outcome demonstrates that our approach obtains competitive performances compared to existing works as depicted in Table III, and it is able to resist to popular distortions which documents often go through.

TABLE III
THE COMPARISON WITH EXISTING APPROACHES

Methods	Document type	Robustness against distortions
Chetan [9]	General content	(1), (2), (5)
Kim [2]	Text content	Rotation, noise insertion, cropping blurring, sharpening, binarization
Varna [6]	Text content	(3), (4)
Palit [1]	Indian text	(2), noise and scaling
Lina [7]	Chinese text	(2), (3), (4), (5), (6)
Our method	General content	(1), (2), (3)

Besides, we also prove the robustness of this work compared to our recent approach [17] by implementing the method [17] on five sample documents, and the results are presented in Table IV. We can see that the accuracy of watermark detection is slightly alleviated. The main factor is due to the extracted watermarking regions less stable than the FCN-based approach. For print-and-scan distortion, both approaches properly work in case of documents printed and scanned at high resolution of 600 dpi.

TABLE IV
THE ACCURATE RATIO OF METHOD [17]

Distortions	Bit Error Rate				
	Doc1	Doc2	Doc3	Doc4	Doc5
Rotation 5° (a)	0	0	0	0	0
Rotation 7°	0.23	0.20	0.22	0.25	0.22
Scaling 0.8	0.17	0	0	0.30	0.14
Scaling 0.9 (b)	0	0	0	0	0
Scaling 1.1 (c)	0	0	0	0	0
Scaling 1.2 (d)	0.11	0	0	0.31	0
(a) + (b)	0	0	0	0	0
(a) + (c)	0	0	0	0	0
(a) + (d)	0.33	0.25	0.28	0.36	0.30
JPEG 60%	0	0	0	0	0
JPEG 50%	0	0	0	0.29	0

V. CONCLUSION

In this work, we propose a blindly invisible watermarking for document images in spatial domain. We transform the problem of document structure segmentation to the problem of watermarking regions detection using FCN and firstly apply this advanced technique in the field of document watermarking for security concern. Compared to existing methods in the same domain, our method achieves high performance in terms of various document contents, robustness, imperceptibility as well as capacity. For future works, we are going to improve the watermarking algorithm such that our approach is able to detect watermark properly in case of complicated distortions as watermarked document either scanned at low resolution or undergone print-photocopy-scan process.

ACKNOWLEDGMENT

This work is supported by the CPER NUMERIC of Nouvelle Aquitaine.

REFERENCES

[1] S. Palit and U. Garain. A novel technique for the watermarking of symbolically compressed documents. In *DIAL'06*, 2006.
[2] Y. W. Kim and O. Seok. Watermarking text document images using edge direction histograms. *Pattern Recogn. Lett.*, 2004.
[3] Y. W. Kim, K.A. Moon, and I. S. Oh. A text watermarking algorithm based on word classification and inter-word space statistics. In *ICDAR '03*, 2003.
[4] Shiyan Hu. Document image watermarking based on weight-invariant partition using support vector machine. In *DAS VI*, 2004.
[5] D. Zou and Y. Q. Shi. Formatted text document data hiding robust to printing, copying and scanning. In *Int. Symposium on Circuits and Systems*, 2005.
[6] A. L. Varna, S. Rane, and A. Vetro. Data hiding in hard-copy text documents robust to print, scan and photocopy operations. In *Int. Conf. on Acoustics, Speech and Signal Processing*, 2009.
[7] S. Guang T. Lina, S. Xingming. Print-scan resilient text image watermarking based on stroke direction modulation for chinese document authentication. *Radioengineering*, 2012.
[8] F. Pingzhi W. Xian K. M. Khurram H. S. Jinn, R. Didi. An adaptive watermarking scheme for e-government document images. *Multimedia Tools Appl.*, 2014.
[9] K.R. Chetan and S. Nirmala. An efficient and secure robust watermarking scheme for document images using integer wavelets and block coding of binary watermarks. *Trans. Info. Security and Appl*, 2015.
[10] B. Noureddine B. Alia, B. Khier. Blind image watermarking technique based on differential embedding in dwt and dct domains. *EURASIP*, 2015.
[11] A. S. Hamid and J. Mansour. Robust watermarking against print and scan attack through efficient modeling algorithm. *Image Commun.*, 2014.
[12] Y. Zolotavkin and M. Juhola. A new two-dimensional quantization method for digital image watermarking. In *Int. Conf. on Advanced Communication Technology*, 2015.
[13] K. Asifullah M. Summuyya. Robust image watermarking technique using triangular regions and zernike moments for quantization based embedding. *Multimed Tools Appl*, 2017.
[14] J. Maedeh, R. Shima, S. S.M.Reza, K. Nader, S. Shahram, N. Kayvan, and S. Shadrokh. Adaptive blind image watermarking using fuzzy inference system based on human visual perception. 2017.
[15] K. Haribabu, G. R. K. S. Subrahmanyam, and D. Mishra. A robust digital image watermarking technique using auto encoder based convolutional neural networks. In *WCI*, 2015.
[16] S. M. Mun, S. H. Nam, H. U. Jang, D. K., and H. K. Lee. A robust blind watermarking using convolutional neural network. *CoRR*, 2017.
[17] V. L. Cu, J. C. Burie, and J. M. Ogier. Stable regions and object fill-based approach for document images watermarking. In *DAS*, 2018.
[18] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. *CoRR*, 2014.
[19] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *CoRR*, 2014.
[20] S. Leutenegger, M. Chli, and R. Y. Siegwart. Brisk: Binary robust invariant scalable keypoints. In *ICCV*, 2011.
[21] A. Apostolos, D. Bridson, C. Papadopoulos, and S. Pletschacher. A realistic dataset for performance evaluation of document layout analysis. ICDAR '09, 2009.
[22] *http://personal.psu.edu/xuy111/projects/cvpr2017_doc.html*.
[23] *http://iapr-tc10.univ-lr.fr/index.php/resources/dataset-and-software/24-resources/databases/296-l3i-textcopies-dataset*.