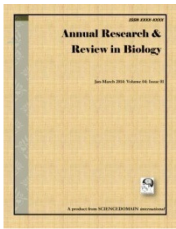

Annual Research & Review in Biology
4(13): 2228-2241, 2014

SCIENCEDOMAIN ***international***
www.sciencedomain.org

Embedding of Pathology Reports in Pathology Images

Hassan Taghipour[1*], Javad Taghipour[2] and Heydar Ali Esmaili[1]

[1]Department of Pathology, Tabriz University of Medical Sciences, Tabriz, Iran.
[2]Department of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran.

Authors' contributions

This work was carried out in collaboration between all authors. Authors HT and JT designed the study, performed the statistical analysis, wrote the protocol, and wrote the first draft of the manuscript. Author HAE managed the analyses of the study and the literature searches. All authors read and approved the final manuscript.

Original Research Article

Received 23rd February 2014
Accepted 23rd March 2014
Published 31st March 2014

ABSTRACT

A pathology report contains Gross or Macroscopic description, Microscopic description and Diagnosis determined by examining cells and tissues under a microscope. In this paper, we propose a method to embed pathology reports in pathology images. In proposed method, adjacent pixel difference is used to embed data. The proposed algorithm is completely appropriate for pathology images. This algorithm has low computational complexity and high embedding capacity. The proposed method is reversible and the original image is restored after extracting data without any loss. The embedding of pathology reports in pathology images can be used in pathology exams, assessment of pathology residents, self-assessment programs, and constructing pathology archives and databases.

Keywords: Pathology images; pathology reports; data hiding; adjacent pixel difference.

**Corresponding author: Email: taghipourh@yahoo.com;*

1. INTRODUCTION

Pathology is a branch of medicine concerned with the etiology, pathogenesis, and pattern of diseases, and also discusses about the changes occurring as a result of diseases in body tissues and organs.

Pathology has a main and important role in diagnosis of different diseases, especially malignant diseases and cancers. Pathologists make diagnosis of diseases based on the gross, microscopic, immunologic and molecular examination of body tissues, and also by laboratory analysis of body fluids and secretions such as blood, urine, stool and CSF using different methods of chemistry, microbiology, hematology and molecular pathology.

In pathology, the tissue removed during a biopsy or surgery must be cut into thin sections, placed on slides, and stained with special dyes, and then it can be examined under a microscope. Nowadays the result of microscopic study of tissues can be stored as digital images, or by using high resolution special scanners, the pathology slides are scanned and stored as digital images.

Every specimen which is examined by a pathologist, has a pathology report. A pathology report is a document that contains the Microscopic description and Diagnosis determined by examining cells and tissues under a microscope. The report may also contain information about the size, shape, color, consistency and appearance of specimen (Gross or Macroscopic description). The pathology report also includes other information, such as patient information (Name, Birthday, Gender, Biopsy date).

For example, in Figs. 1 and 2, we present two pathology images with their “Microscopic description” and “Diagnosis”:

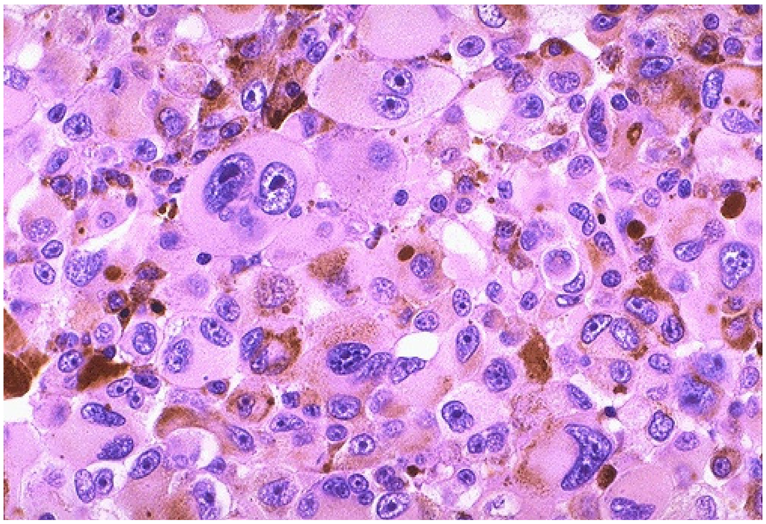

Fig. 1. Malignant Melanoma

Microscopic description: This is biopsy of skin lesion, composed of large polygonal cells with marked pleomorphic nuclei, which contain prominent nucleoli. The neoplasm is making melanin pigment.

Diagnosis: Malignant melanoma

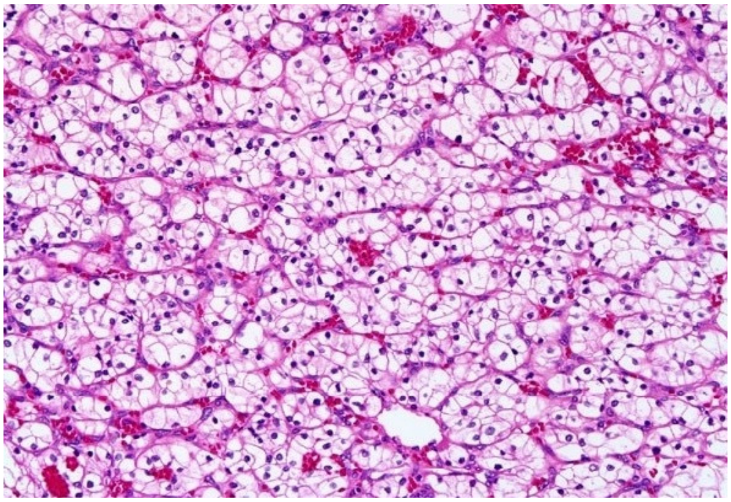

Fig. 2. Clear cell carcinoma of kidney

Microscopic description: This is biopsy of kidney lesion, which is composed of large polygonal cells with clear cytoplasm and distinct cell membranes. The nuclei are small, round to slightly oval, and do not show any morphologic atypia, conforming the nuclear grade ¼. The cells arranged in nests with intervening blood vessels.

Diagnosis: Clear cell carcinoma of kidney.

Data embedding or watermarking is a technique to embed invisible information bits into a host media. Data embedding is commonly used in digital images. It is desired that after extracting embedded data, the original image can be restored without any loss. These methods of data embedding are called lossless or reversible data embedding.

Several methods have been proposed for reversible data embedding in digital images. In [1,2], a reversible data embedding technique has been proposed by using difference expansion. Sliding window has been used to embed data in [3]. In [4-6], wavelet is used for data hiding. Wavelet based methods have high computational complexity. Ni et al. [7] used the maximum and minimum points of the image histogram for embedding data. Ni et al.'s method is reversible and has low complexity. Reversible data scheme based on adjacent pixel difference has been proposed in [8] to improve Ni method. The adjacent pixel difference method has low computational complexity and high embedding capacity.

In this paper, we want to embed pathology reports (including: patient information, Gross description, Microscopic description and Diagnosis) in corresponding pathology images without any significant changes in the quality of images. The embedding of pathology reports in pathology images can be used in pathology exams, assessment of pathology residents, self-assessment programs, and constructing pathology archives and databases. The proposed method is based on the adjacent pixel difference. The pathology images are color images and have special characteristics. The proposed algorithm is completely suitable for pathology images.

The rest of the paper is organized as follows. In section 2, the method of adjacent pixel difference is presented. The proposed method for embedding data in pathology images is presented in section 3. The experimental results are discussed in section 4. Finally, concluding remarks are presented in section 5.

2. THE METHOD OF ADJACENT PIXEL DIFFERENCE

In the method of adjacent pixel difference [8], the cover image I is supposed to be an 8-bit grayscale image with size $M \times N$. The adjacent pixel difference in row direction is defined as:

$$d(i, j) = I(i, j) - I(i, j+1); \quad 1 \le i \le M, \quad 1 \le j \le N-1, \tag{1}$$

Where $I(i, j)$ denotes the pixel value at location (i, j) . In Fig. 3, the histogram of pixel difference of a typical image is shown. Embedding procedure is based on making null in histogram of difference. The difference, whose absolute value equals to δ , is used for embedding data. To make null in difference histogram, the differences, whose value are greater than δ , are added by 1, and those with value less than $-\delta$, are subtracted by 1. This process can be described as:

$$d'(i, j) = \begin{cases} d(i, j)+1, & d(i, j) > \delta \\ d(i, j)-1, & d(i, j) < -\delta . \\ d(i, j), & otherwise \end{cases} \tag{2}$$

After this process, there is no pixel difference with value of $(\delta+1)$ and $-(\delta+1)$. As shown in Fig. 4, there are two nulls in the histogram. These nulls are utilized to embed data. If the embedding data bit is "1", d' with value of δ will increase 1 or d' with value of $-\delta$ will decrease 1. If the embedding data bit is "0", the difference value of δ or $-\delta$ remains intact. The new difference value is obtained as:

$$d^*(i, j) = \begin{cases} d'(i, j)+1, & d'(i, j) = \delta \ and \ bit = 1 \\ d'(i, j)-1, & d'(i, j) = -\delta \ and \ bit = 1. \\ d'(i, j), & otherwise \end{cases} \tag{3}$$

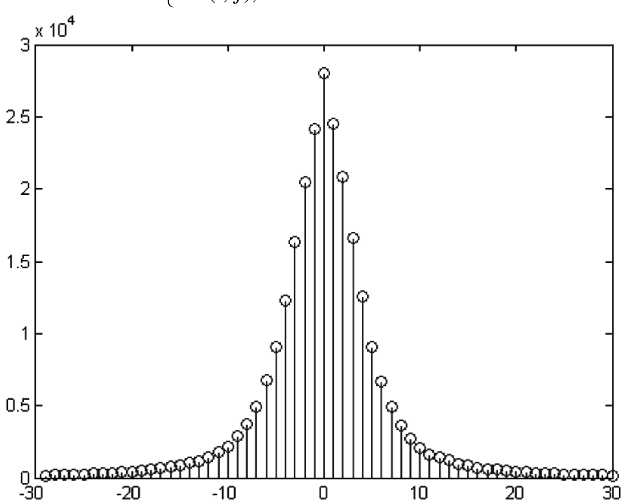

Fig. 3. Histogram of pixel difference of a typical image [8]

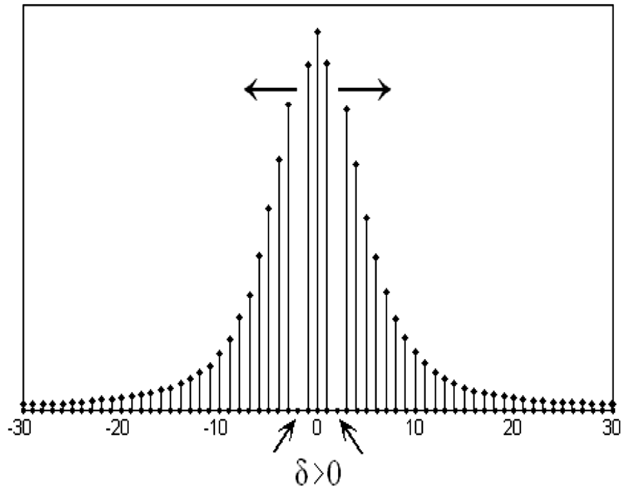

Fig. 4. Histogram of pixel difference after making nulls [8]

The image after embedding data is called stego image or watermarked image. The stego image I^* can be calculated as:

$$I^*(i,j) = I(i,j+1) + d^*(i,j); \quad 1 \le i \le M, \quad 1 \le j \le N-1. \tag{4}$$

In this process, the last column of the original image $(j = N)$ remains unchanged, therefore:

$$I^*(i,j) = I(i,j); \quad j = N. \tag{5}$$

The capacity of this method is the number of differences with value of δ or $-\delta$.

The embedding data can be extracted from stegoimage I^*. Extracting data and restoring original image from stego image I^* is done from right to left and started from last column of the stego image. According to equation (5) the last column of stego image and original image is the same. At first, the last column of original image is obtained as:

$$I(i,N) = I^*(i,N) \tag{6}$$

Then the pixel difference of the stego image is calculated as:

$$d^*(i,j) = I^*(i,j) - I(i,j+1). \tag{7}$$

According to the embedding algorithm, the differences with values δ or $-\delta$ show that the embedded bit is "0" and the differences with values $\delta+1$ or $-(\delta+1)$ show that the embedded bit is "1". So, if $d^*(i,j)$ equals to δ or $-\delta$, a bit "0" will be extracted and if

$d^*(i,j)$ equals to $\delta+1$ or $-(\delta+1)$, a bit "1" will be extracted. The original pixel difference value can be calculated as:

$$d(i,j)=\begin{cases} d^*(i,j)-1, & d^*(i,j)\geq\delta+1 \\ d^*(i,j)+1, & d^*(i,j)\leq-(\delta+1). \\ d^*(i,j), & otherwise \end{cases} \tag{8}$$

Then the original pixel value can be restored as:

$$I(i,j)=I^*(i,j+1)+d(i,j) \tag{9}$$

To extract all of the embedded data and restore original image, the equations (7) to (9) are repeated until j=1.

In embedding procedure, pixel values may be increased or decreased. For pixel values of 0 or 255 in an 8-bit grayscale image, underflow and overflow will occur. To prevent underflow and overflow the pixel values of 0 and 255 are reset to 1 and 254 respectively. The locations of these pixels are embedded with data to restore the original image without any loss [8].

3. PROPOSED ALGORITHM FOR PATHOLOGY IMAGES

In this section, we propose a method to hide data in pathology images by using adjacent pixel difference. Proposed algorithm is completely appropriate for data hiding in pathology images. Proposed algorithm is described in three subsections as follow.

3.1 Algorithm for Color Images

The pathology images are color images. In Figs. 1 and 2, two typical pathology images are illustrated. The RGB (red, green, blue) color model is the most commonly used model for color images. Images represented in RGB color model consist of three component images, one for each primary color. For example, a 24-bit color image consist red, green and blue components, which each component is an 8-bit image [9].

For each of these components, we can embed data by using pixel differences. The pixel differences of each component can be calculated as:

$$\begin{cases} d_r(i,j)=I_r(i,j)-I_r(i,j+1); & 1\leq i\leq M, \quad 1\leq j\leq N-1, \\ d_g(i,j)=I_g(i,j)-I_g(i,j+1); & 1\leq i\leq M, \quad 1\leq j\leq N-1, \\ d_b(i,j)=I_b(i,j)-I_b(i,j+1); & 1\leq i\leq M, \quad 1\leq j\leq N-1, \end{cases} \tag{10}$$

Where I_r, I_g and I_b are the red, green and blue components of image I respectively, d_r, d_g and d_b are the pixel difference of these components. The embedding algorithm is separately applied for each of these difference values.

3.2 Proposed Algorithm to Prevent Overflow

Almost all of pathology images are bright images. In these images, there are many pixels with high values. In Fig. 5, the histograms of red, green and blue components of Fig. 1 are shown. The number of pixels with high values (for example 254 or 255) is numerous in each of these components, and the number of pixels with low values is few or zero. Therefore, in these images the overflow will occur during data embedding. Because the number of pixels with value of 255 is high, it is impossible that the location of these pixels is embedded with data for restoring original image.

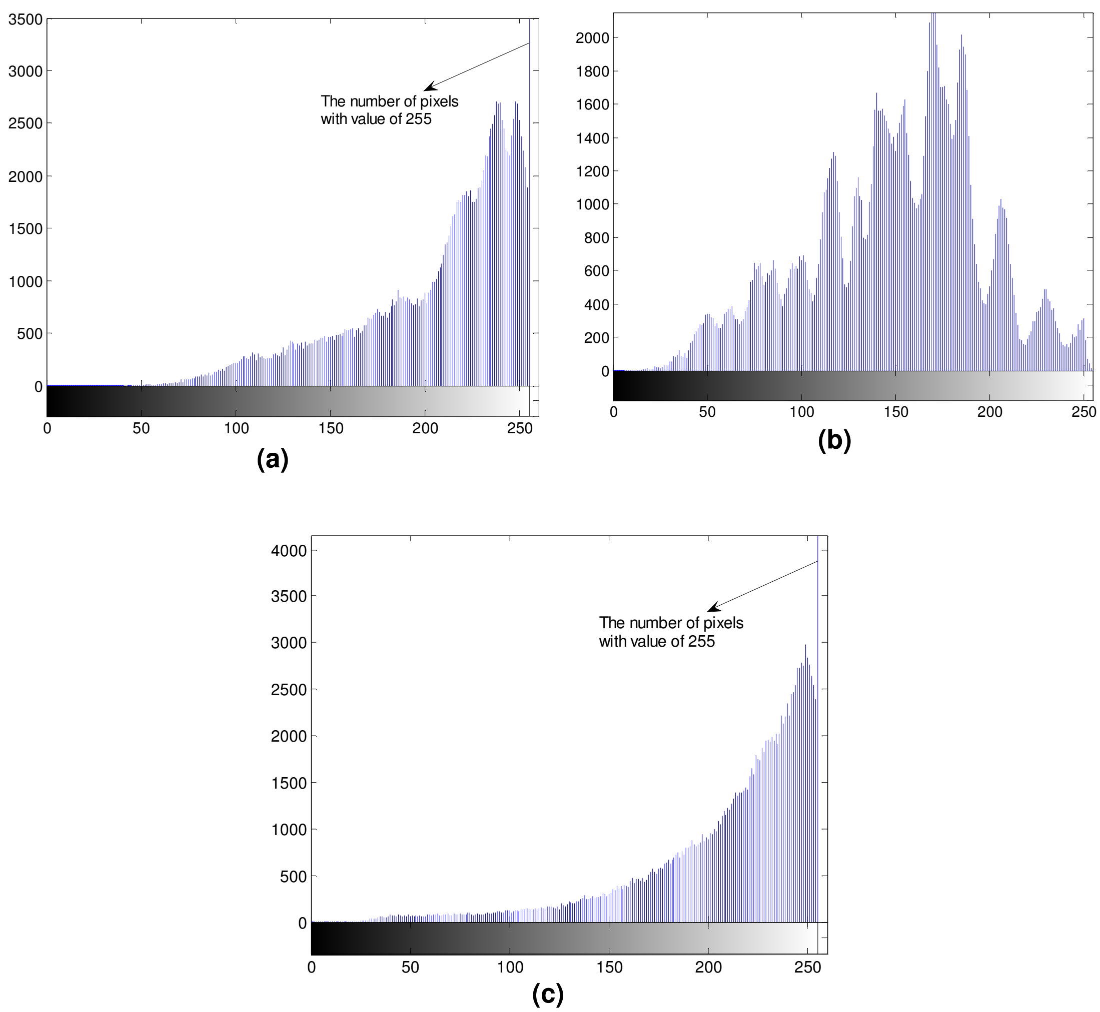

Fig. 5. The histograms of (a) red, (b) green and (c) blue components of Fig. 1

In embedding procedure, if the pixel values are not increased, the overflow will not occur. Therefore, proposed algorithm is based on decreasing pixel values to prevent overflow. In order to decrease pixel value in embedding procedure, the histogram of difference just is shifted to left side. For example, suppose that the differences $\delta = a_1, a_2, a_3$ have been chosen to embed data. Based on these selected differences, the histogram of difference is shifted to left side as:

$$d'(i,j)=\begin{cases} d(i,j)-3, & d(i,j)<a_1 \\ d(i,j)-2, & a_1 \le d(i,j)<a_2 \\ d(i,j)-1, & a_2 \le d(i,j)<a_3 \\ d(i,j), & otherwise \end{cases}, \quad a_1<a_2<a_3. \tag{11}$$

After this process, there is no pixel difference with value of a_1-3, a_2-2 and a_3-1. These differences are used to embed data. If the embedding data bit is "1", d' with value of a_1-2, a_2-1 or a_3 will decrease 1. If the embedding data bit is "0", the difference value of a_1-2, a_2-1 or a_3 remains intact. The new difference value is obtained as:

$$d^*(i,j)=\begin{cases} d'(i,j)-1, & d'(i,j)=a_1-2 \ \ and \ \ bit=1 \\ d'(i,j)-1, & d'(i,j)=a_2-1 \ \ and \ \ bit=1 \\ d'(i,j)-1, & d'(i,j)=a_3 \ \ and \ \ bit=1 \\ d'(i,j), & otherwise \end{cases}. \tag{12}$$

The stego image I^* can be calculated from equation (4). Embedded data can be extracted from stego image. The pixel difference d^* can be obtained from equation (7). If $d^*(i,j)$ equals to a_1-2, a_2-1 or a_3, a bit "0" will be extracted and if $d^*(i,j)$ equals to a_1-3, a_2-2 or a_3-1, a bit "1" will be extracted. The original pixel value can be obtained as:

$$d(i,j)=\begin{cases} d^*(i,j)+1, & a_2-1 \le d^*(i,j) \le a_3-1 \\ d^*(i,j)+2, & a_1-2 \le d^*(i,j) \le a_2-2 \\ d^*(i,j)+3, & d^*(i,j) \le a_1-3 \\ d^*(i,j), & otherwise \end{cases}. \tag{13}$$

Finally, the original pixel value can be restored from equation (9).

3.3 Proposed Algorithm for Increasing Embedding Capacity

In a digital image, each pixel has eight adjacent pixels. For a pixel in location (i,j), eight adjacent pixels are shown in Fig. 6. All of these adjacent pixels can be used for embedding data to increase the embedding capacity. The pixel difference values for these adjacent pixels can be calculated as:

$$\begin{cases} d_1(i,j) = I(i,j) - I(i-1,j-1); & 2 \le i \le M, \quad 2 \le j \le N, \\ d_2(i,j) = I(i,j) - I(i-1,j); & 2 \le i \le M, \quad 1 \le j \le N, \\ d_3(i,j) = I(i,j) - I(i-1,j+1); & 2 \le i \le M, \quad 1 \le j \le N-1, \\ d_4(i,j) = I(i,j) - I(i,j+1); & 1 \le i \le M, \quad 1 \le j \le N-1, \\ d_5(i,j) = I(i,j) - I(i+1,j+1); & 1 \le i \le M-1, \quad 1 \le j \le N-1, \\ d_6(i,j) = I(i,j) - I(i+1,j); & 1 \le i \le M-1, \quad 1 \le j \le N, \\ d_7(i,j) = I(i,j) - I(i+1,j-1); & 1 \le i \le M-1, \quad 2 \le j \le N, \\ d_8(i,j) = I(i,j) - I(i,j-1); & 1 \le i \le M, \quad 2 \le j \le N, \end{cases} \tag{14}$$

Where $d_k(i, j)$ is the difference value of *kth* adjacent pixel. We can embed data in each of these differences separately to increase the embedding capacity.

1	2	3
8	I(i,j)	4
7	6	5

Fig. 6. The adjacent pixels of I (i,j)

4. EXPERIMENTAL RESULTS

In this section, three pathology color images are used to analysis the performance of proposed method. These images are shown in Figs. 1, 7 and 8. The size of these images is 330×504. The embedding capacity and peak signal to noise ratio (PSNR) are evaluated as two important factors for performance. The embedding capacity is measured in bit per pixel (bpp). The PSNR is used as the distortion measurement. The PSNR of the stego image is calculated as:

$$PSNR = 10 \log_{10} \left(\frac{Max_I^2}{MSE} \right), \tag{15}$$

Where Max_I is the maximum pixel value of the image and MSE (Mean Square Error) is calculated as:

$$MSE = \frac{1}{M \times N} \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} \left(I(i,j) - I^*(i,j) \right)^2 \tag{16}$$

Where I and I^* are the original and stego images respectively with size $M \times N$. For RGB color images, MSE is the average of the MSE values of three components. The PSNR shows the quality of the stego image.

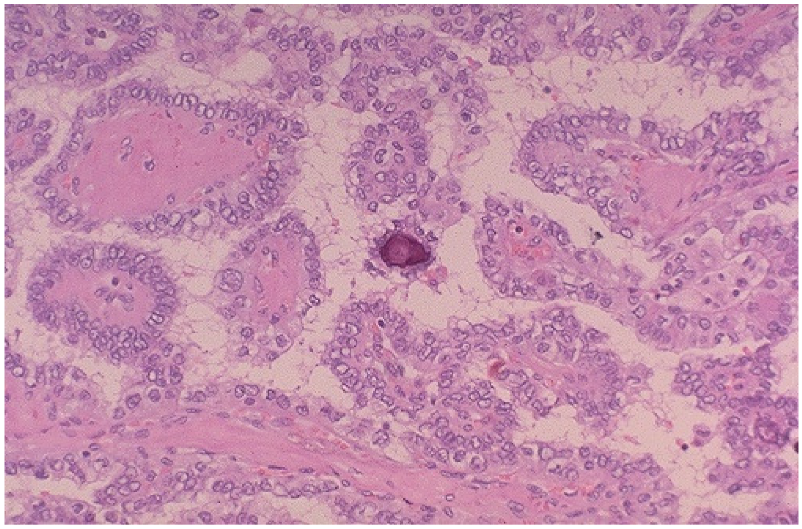

Fig. 7. Papillary carcinoma of Thyroid

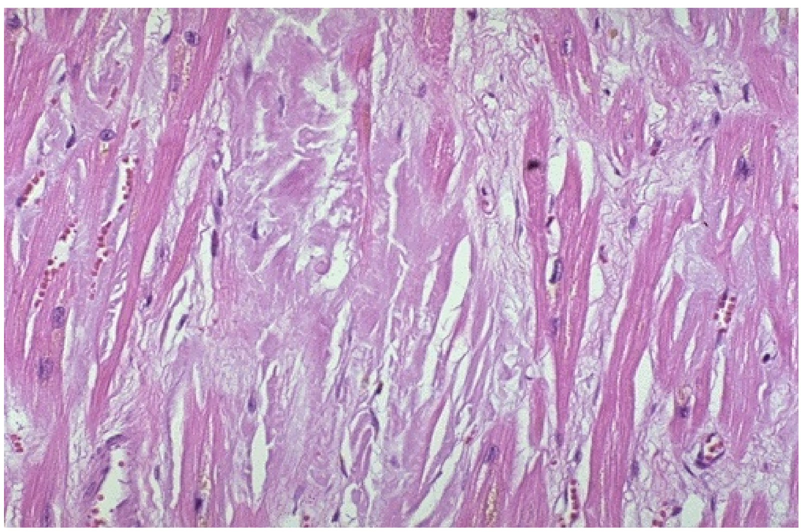

Fig. 8. Amyloidosis of Myocardium

In Fig. 9, embedding capacity is plotted against the PSNR for images of Figs. 1, 7 and 8. In this figure, adjacent pixels 2, 4, 6 and 8 have been used for embedding. It is clear that the PSNR decreases when the embedding capacity increases. For embedding capacity 0.8bpp, the PSNR value remains greater than 30dB.

Some pathology images are not bright enough like the images that are shown in Figs. 2 and 10. The histograms of pixel value for green components of these figures are shown in Fig. 11. For green component, there are a lot of pixels with value of 0. In embedding procedure, underflow will occur for green component of these images. We studied more than 30 different pathology images, and in some of them, only the green component had this problem. To avoid underflow, we do not embed any data in green components of these images based on the algorithm that has been explained in subsection 3.2.

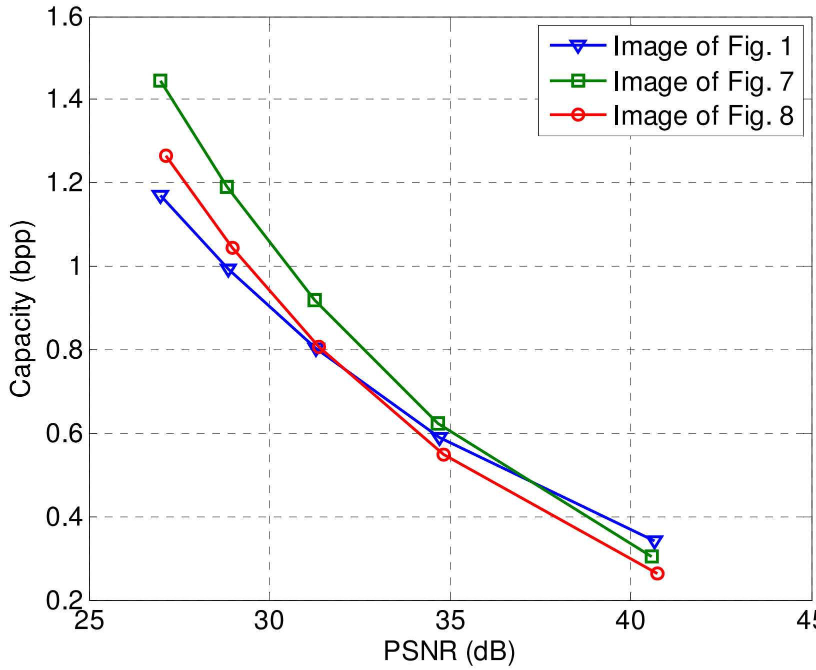

Fig.9. Performance of the proposed algorithm for three pathology images

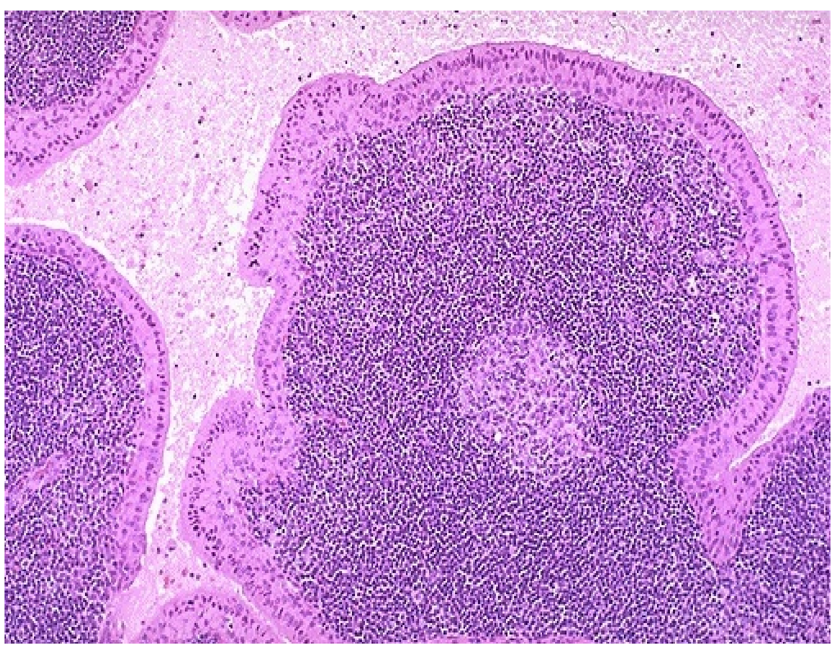

Fig. 10. Warthin's tumor

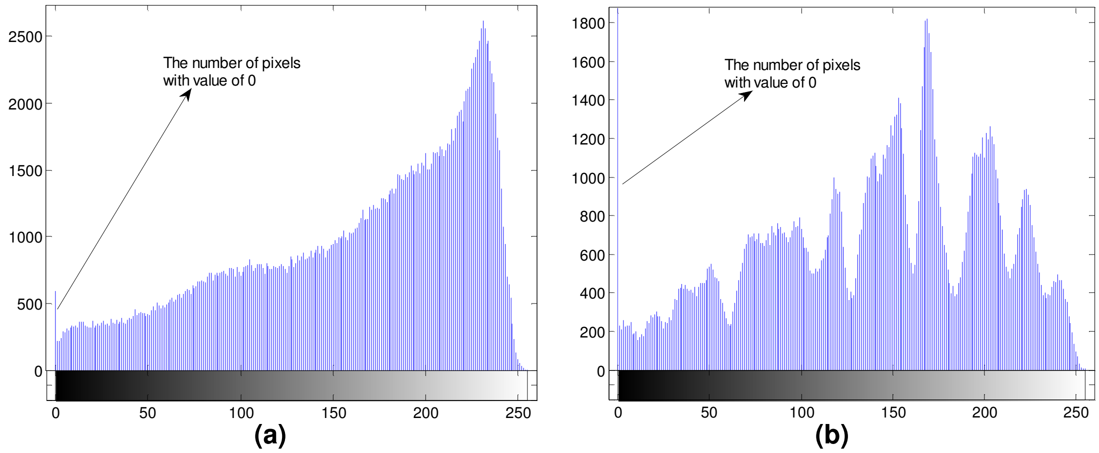

Fig. 11. The histogram of green component, (a) Fig. 2 (b) Fig. 10

If the number of pixels with high values is few or zero in green component of these images, we can modify the embedding algorithm for green component. For embedding data in green component of these images, the equations (11) and (12) can be modified as:

$$d'(i,j)=\begin{cases} d(i,j)+3, & d(i,j)>a_3 \\ d(i,j)+2, & a_2<d(i,j)\le a_3 \\ d(i,j)+1, & a_1<d(i,j)\le a_2 \\ d(i,j), & otherwise \end{cases}, \quad a_1<a_2<a_3. \tag{17}$$

$$d^*(i,j)=\begin{cases} d'(i,j)+1, & d'(i,j)=a_1 \ and \ bit=1 \\ d'(i,j)+1, & d'(i,j)=a_2+1 \ and \ bit=1 \\ d'(i,j)+1, & d'(i,j)=a_3+2 \ and \ bit=1 \\ d'(i,j), & otherwise \end{cases}. \tag{18}$$

Therefore, in embedding procedure the pixel values are increased and the underflow will not occur.

Fig. 12 shows the stego image of Fig. 1. In this image, 195104 bits (1.17bpp) are embedded and the PSNR is 27dB, but we easily can diagnose it as Malignant Melanoma, however after extracting data the original image will be restored.

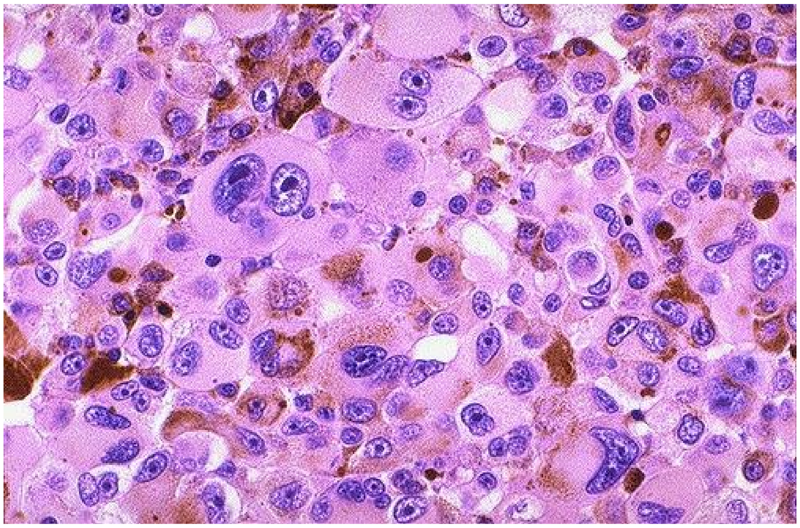

Fig. 12. Stego image of Fig. 1

5. CONCLUSION

In this paper, we applied the pixel value differencing method to embed pathology reports in pathology images. Proposed algorithm was completely appropriate for pathology images. In proposed method, after extracting data, the original image can be restored without any loss. Experimental results showed that the PSNR or quality of the stego image decreases by increasing the embedding capacity. In a typical pathology image with size 330×504, about 194000 bits (1.17bpp) was embedded and the PSNR was about 27dB. The proposed method is reversible and the original image is restored completely, however the pathology diagnosis can be performed by using stego image without any problem. The embedding of pathology reports in pathology images can be used in pathology exams, assessment of pathology residents, self-assessment programs, and constructing pathology archives and databases.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Tian J. "Reversible Data Embedding Using a Difference Expansion," IEEE Transactions on Circuits and Systems for Video Technology. 2003;(13)8:890-896,
2. Alattar AM. "Reversible Watermark Using the Difference Expansion of a Generalized Integer Transform," IEEE Transactions on Image Processing. 2004;(13)8:1147-1156,
3. Lee CF, Chen HL. "Reversible Data Embedding Using the Sliding Window", IEEE International Conference on Intelligent Information Hiding and Multimedia Signal Processing. 2008;118–1121.
4. Agreste S, Andaloro G, Prestipino D, Puccio L. “An image adaptive, wavelet-based watermarking of digital images,” Journal of Computational and Applied Mathematics. 2007;(210)1-2:13–21,

5. Al-Otum HM, Samara NA. “A robust blind color image watermarking based on wavelet-tree bit host difference selection,” Signal Processing. 2010;(90)8:2498–2512.
6. Liu G, Liu H, Kadir A. “Wavelet-Based Color Pathological Image Watermark through Dynamically Adjusting the Embedding Intensity,” Computational and Mathematical Methods in Medicine. 2012;1-10.
7. Ni Z, Shi YQ, Ansari N, Su W. "Reversible Data Hiding" IEEE Transactions on Circuits and Systems for Video Technology. 2006(16)3:354-362.
8. Li Z, Chen X, Pan X, Zeng X. "Lossless Data Hiding Scheme Based on Adjacent Pixel Difference," IEEE computer society, International Conference on Computer Engineering and Technology. 2009;588-592.
9. Gonzalez RC, Woods RE. Digital Image Processing. Third Edition, Prentice Hall; 2008.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history.php?iid=463&id=32&aid=4200