Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 46 (2015) 135 – 142

Procedia Computer Science

International Conference on Information and Communication Technologies (ICICT 2014)

An Improved Automatic Detection of True Comets for DNA Damage Analysis

Sreelatha G.[a,*], Aparna Muraleedharan[b], Parkash Chand[b], Ravi Philip Rajkumar[c], P. S. Sathidevi[a]

[a]*Dept. of Electronics & Communication, National Institute of Technology Calicut, Kozhikode - 673 601, India.*
[b]*Dept. of Anatomy, Jawaharlal Institute of Postgraduate Medical Education & Research, Puducherry - 605 006, India.*
[c]*Dept. of Psychiatry, Jawaharlal Institute of Postgraduate Medical Education & Research, Puducherry - 605 006, India.*

Abstract

Comet assay is a well known method for evaluating DNA damage in individual cells. Silver stained comet assay images are employed in this study as they are widely used in clinical applications. When compared with fluorescent images silver stained images require additional pre-processing stages to effectively identify the region of interest. An improved method for effective automatic identification of true comets from the noisy silver stained images is proposed in this paper. The performance of the proposed method is compared with the recent related methods and shows better results in terms of %TP, %FN, %FP, Positive Predictive Value and sensitivity.

Peer-review under responsibility of organizing committee of the International Conference on Information and Communication Technologies (ICICT 2014)

Keywords: DNA damage ; Comet detection ; Silver stained comet assay images ; Comet assay

1. Introduction

The comet assay or single cell gel electrophoresis is a well known and widely used method for DNA damage analysis because of its simplicity, low cost, the collection of data at individual cell level allowing robust type of statistical analysis, the requirement of only small number of cells per sample and high sensitivity for detecting low levels of DNA damage. It was first introduced in 1984 by Ostling and Johanson as a method to measure DNA single strand breaks that caused relaxation of DNA supercoils under neutral conditions[1]. Later, a modified version was introduced by Singh *et al.* in 1988, which used alkaline conditions that improved its specificity and reproducibility[2]. This method is capable of measuring single-strand or double-strand DNA breaks, DNA cross-links, alkali labile sites, base/base-pair damages and apoptotic nuclei in the cell. It has widespread applications in the area of testing novel chemicals for genotoxicity, monitoring environmental contamination with genotoxins, human biomonitoring and molecular epidemiology, diagnosis of genetic disorders and fundamental research in DNA damage and repair.

* Corresponding author. Tel.: +91-949-629-1852 ; fax: +0-000-000-0000.
E-mail address: sreelatha_ pec12@nitc.ac.in

Peer-review under responsibility of organizing committee of the International Conference on Information and Communication Technologies (ICICT 2014)
doi:10.1016/j.procs.2015.02.004

DNA damage analysis provides valuable information regarding early biological effects of occupational exposure to hazardous chemicals and important information about the nature and type of cancer, which could be used by oncologists for treatment planning and determining the best course of intervention. There are significant studies going on relating to the effect of smoking, eating habits, use of mobile phones, use of cosmetics etc. on DNA damage[3]. Thus it is evident that the study and analysis of DNA damage assumes great importance in clinical research and hence there is a clear demand for an accurate, fast and sensitive method to detect DNA damage[4].

2. Comet assay methodology

The comet assay protocol was developed by Olive *et al.* in 2006[5]. The basic steps involved in this procedure are sample collection, lymphocyte separation, slide preparation, cell lysis, alkali treatment, electrophoresis, neutralization, DNA staining, comet visualization and comet scoring. A detailed description of this methodology is available in[6,7,8]. Fig. 1 shows the overview of comet assay methodology.

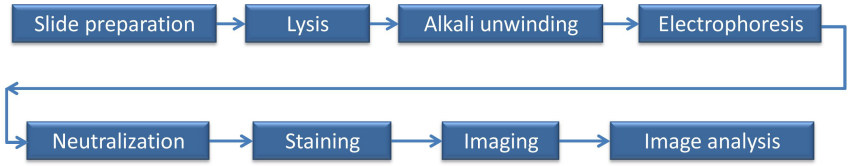

Fig. 1. Overview of Comet Assay Methodology

During electrophoresis the damaged cells acquire the shape of a comet, hence the cells are called comets and the procedure comet assay. These comets can be visualized using fluorescent staining method or silver staining method. In this study, we have considered Silver stained comet assay images because in clinical applications, they are better preferred. This is because, for fluorescent staining and photomicrography, high quality fluorescence microscope is required. Further, the slides cannot be stored for a long period of time and they should be photographed and analysed immediately[9]. The advantages of Silver staining method is that it is inexpensive, slides can be preserved for a long time, it is less hazardous and the analysis can be carried out using a simple light microscope, which is available in all laboratories. But the disadvantage when compared with fluorescent stained images is that silver stained images are very noisy. To have a better visualization of the above facts some sample images are given in Fig. 2. Therefore, detecting the true comets from the noisy silver stained images is a challenging task.

Fig. 2 (a) & (b) show some of the silver stained comet assay images viewed under bright-field light microscope and Fig. 2 (c) shows fluorescent stained comet assay images. While analysing Fig. 2, it is clear that for fluorescent images the background is clear, but for silver stained images the background is not homogeneous and also the tail region (refer Fig. 2 (b)) is merged with background so the detection of comets or proper selection of ROI (Region of Interest) without loss of tail information is a challenging task. Therefore, detecting comets from silver stained images require extra pre processing stages. The major tools available today for comet assay image analysis are of two types - semi automated and fully automated. Semi automated tools support the analysis of both silver stained and fluorescent stained comet assay images, because selection of comets is manual with some markers[10,11,12]. Most of the fully automated software tools are developed for fluorescent stained images[13,14,15,16]. The OpenComet software package support both types of images[17].

3. Comet detection

Comet assay image analysis has three stages - Comet detection, Comet segmentation and Comet quantification. This paper proposes an improved method for comet detection stage. The major steps involved in comet detection stage are described in Fig. 3.

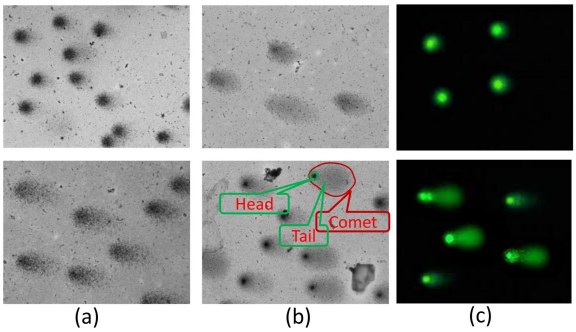

Fig. 2. (a) & (b) Examples of silver stained comet assay images ; (c) Examples of fluorescent stained comet assay images

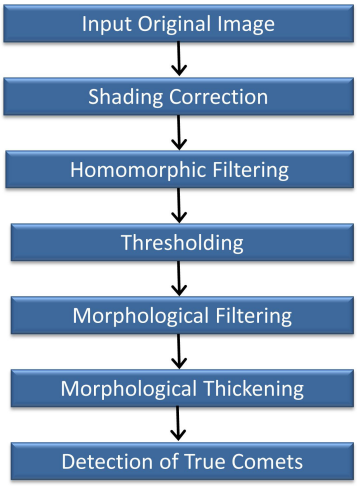

Fig. 3. Block level representation of automatic comet detection stage

3.1. Shading correction

Microscopic images undergo shading effect due to the uneven illumination of the field while capturing the image. Therefore shading correction is done as an initial step in the proposed method. Gray scale morphological bottom-hat or top-hat transformation can be used for shading correction [18]. Equation (1) indicates bottom-hat transformation, where $I_{Bo}(x, y)$ is the bottom hat filtered image, $F(x,y)$ is the input image, b is the structuring element and '•' represents gray scale closing operation. Shading corrected image $I_s(x, y)$ is obtained by subtracting $I_{Bo}(x, y)$ from $F(x,y)$. $I_s(x, y)$ is then kept as the reference image for further analysis.

$$I_{Bo}(x, y) = [F(x, y) \bullet b] - F(x, y) \tag{1}$$

$$I_s(x, y) = F(x, y) - [F(x, y) \bullet b] \tag{2}$$

3.2. Homomorphic filtering

In silver stained comet assay images the tail region of the comets will be merged with the background. Therefore in order to improve the appearance of the comets in the image, homomorphic filtering technique is used which simul-

taneously compresses the intensity range and enhances the contrast. Initially the shading corrected image is converted to log domain. Then FFT of the log transformed image is computed. Then Gaussian high-pass filtering is done with a standard deviation of 0.2. Further inverse FFT is computed and finally the homomorphic filtered image is obtained by applying the exponential function to the resultant image.

3.3. Thresholding

After shading correction and image enhancement, the next stage is thresholding. Otsu's thresholding method[19] is used to get a binary image. The binary image thus obtained will contain lot of noises which have to be removed. These noises are removed using morphological filtering operations.

3.4. Morphological filtering operations

The first step in this process is morphological closing operation which effectively smoothen the contours, eliminates small holes and fills smaller gaps in the object boundaries. Further, morphological filling operation is carried out to fill further holes (which can be found in the case of highly damaged cells) present in the object. Then opening operation is done to eliminate noise in the background. A disk of radius 4 is used as structuring element for all morphological operations.

3.5. Morphological thickening operation

After morphological filtering operation thickening operation is carried out. This is done to ensure that the selected ROIs will not lose any tail information in the background. This operation helps to grow the selected ROIs and hence gives better result in comet parameter quantification.

3.6. Detection of true comets

The final stage is to detect the true comets from the selected ROIs. All the objects may not be true comets, it may contain overlapping comets, artifacts and boundary objects. In this step we need to eliminate overlapping comets, reject gel or air artifacts and remove any objects near the image boundary.

The first step in this process is removal of boundary objects. Boundary objects are rejected by removing the 8 connected high intensity objects at the image border. Once boundary objects are removed, the next step is to eliminate overlapping comets. For this borders of the selected objects are projected onto the reference image $I_s(x, y)$. Then individual comets are cropped out from $I_s(x, y)$. The comets should be orientation corrected to place them in horizontal direction to effectively quantify the comet parameters, especially comet length and similar parameters. After orientation correction, overlapping comets are removed by thresholding each selected object at 80% (head region will have an intensity in the range of 100-80%) of the maximum intensity. If overlapping comets are there then there will be multiple objects. And hence such objects with multiple elements can be rejected. For heavily damaged cells, thresholding can generate multiple regions for a single cell itself. So the algorithm will reject these type of non overlapping comets. In order to solve this problem morphological closing and filling operations are carried out. Refer Fig. 4 (a) & (b).

The air or gel artifacts can be removed by analysing the profile of each object. The profile of n^{th} comet is computed as in Equation (3), where m is the number of pixels along the y axis of the n^{th} object at x position, $I_n(x, y)$ is the intensity level at coordinates (x, y) of the n^{th} object, y_1 and y_m are the lowest and highest y coordinate of the n^{th} object at x position. Typically, artifacts are characterized by low and homogeneous intensity and, therefore, the n^{th} object shows a low value of $d_n = max[P_n(x)] - average[P_n(x)]$.

$$P_n(x) = \frac{1}{m} \sum_{y=y_1}^{y_m} I_n(x, y) \tag{3}$$

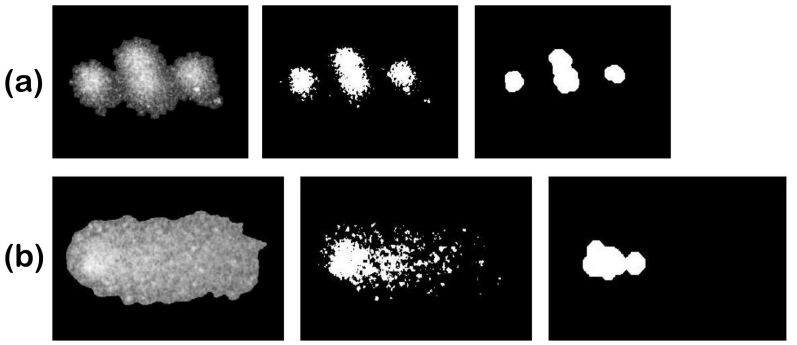

Fig. 4. Elimination of overlapping comets by thresholding at 80% of the maximum intensity : (a) Moderately damaged overlapping cells ; (b) Heavily damaged cell

Considering N objects within an image, the n^{th} object is rejected as an artifact if the following condition is satisfied.

$$d_n < T_d \frac{1}{N} \sum_{j=1}^{N} d_j \tag{4}$$

The value of T_d was experimentally determined and was fixed at 0.70. Sometimes, due to noise in the background, unwanted areas may get selected. These areas are removed by estimating the average size of the comets present in the image. Areas with less than half of the average area and greater than twice the average area of the comet are rejected. At this stage only true comets are left out.

4. Result & performance analysis

After obtaining written and informed consent, blood samples were drawn from a peripheral vein of confirmed cases of Schizophrenia between the age group of 18 to 65 years admitted in Psychiatry ward of JIPMER, and lymphocytes were separated by centrifugation. The slides stained with silver nitrate were viewed under a bright-field light microscope (Olympus BX43, 20x magnification) and the images were captured using CCD camera. These images were used for comet assay image analysis to quantify DNA damage. The proposed method was tested for different cases (patients having different levels of Schizophrenia). The performance of the proposed method was compared with Sansone's *et al.*[14] method and Gyori's *et al.*[17] method. Sansone's *et al.* work is customised for fluorescent stained images. This method could not effectively select the entire comet without loss of tail information in the case of silver stained comet assay images. Refer Fig. 5(b). Gyori's *et al.* developed an open source software called OpenComet (available at *http://www.opencomet.org*), where comet detection is based on geometric shape attributes. Since comet detection is based on shape attributes some of the noises having similar structure as that of comets may be detected as true comets in the detection stage. Refer Fig. 5. The number of true comets present in the image (refer Fig. 5(a)) is 9 but OpenComet software (refer Fig. 5(c)) selected 14 comets as true comets. Five of them are incorrectly detected as comets. Fig. 5 shows the result of comet detection stage of the proposed method and that of the other two methods.

Performance of the detection algorithm is validated by clinical experts and also by a confusion matrix. A clinical expert selected the acceptable comets from our data set. The total number of comets accepted by the operator is the True Number of Comets (TNC) in the images. A comet is considered to be correctly detected if it is selected by the operator also. In addition to True Positive (TP : Number of correctly detected comets by the proposed method), False Negative (FN: Number of comets not detected by the proposed method) and False Positive (FP: Number of falsely detected comets by the proposed method), the following indices of performance are also calculated.

- PPV (Positive Predictive Value):

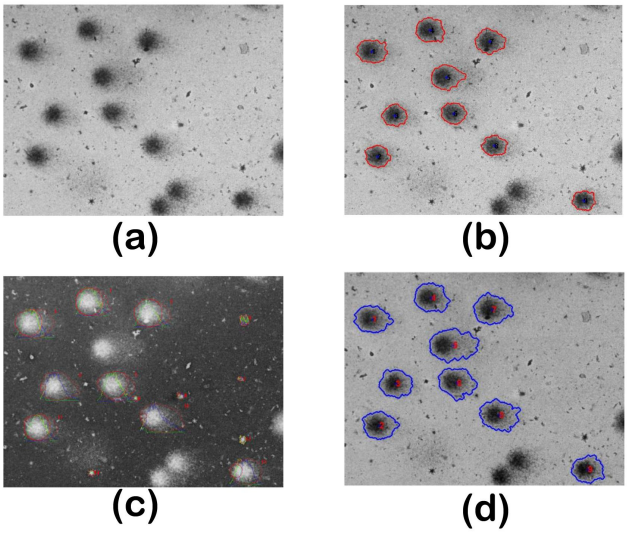

Fig. 5. (a) Original Image ; (b) Result of Sansone's *et al.* method [14] ; (c) Result of Gyori's *et al.* method [17] ; (d) Result of the proposed method

$$PPV = \frac{TP}{TP + FP} \tag{5}$$

- Sensitivity :

$$Sensitivity = \frac{TP}{TP + FN} \tag{6}$$

%TP, %FN and %FP were calculated as percentage of TNC. Sansone's *et al.* method was implemented and its performance was tested on our data set. Similarly the data sets were tested with OpenComet software also. For this study we have considered five different cases and altogether there are 90 comet assay images. The total number of comets present in these images is 429, among which 335 true comets are identified by the clinical expert. The results are tabulated in Table 1, 2 and 3. Table 4 summarises the result of the three methods. When compared with the related recent methods, the proposed method gives better results in terms of %TP, %FN, %FP, PPV and sensitivity.

Table 1. Performance analysis of Sansone's *et al.* method [14]

Cases	Total No. of Comets	TNC	Sansone's *et al.* method [14]				
			%TP	%FN	%FP	PPV	Sensitivity
Case 1	60	43	75	25	43.182	0.635	75
Case 2	128	104	90.722	9.278	14.433	0.863	90.722
Case 3	85	76	82.278	17.722	1.266	0.985	82.278
Case 4	50	38	70.732	29.268	0	1	70.732
Case 5	106	74	76.056	23.944	11.268	0.871	76.056
Average	429	335	**78.96**	**21.04**	**14.03**	**0.871**	**78.96**

Table 2. Performance analysis of Gyori's *et al.* method [17]

Cases	Total No. of Comets	TNC	OpenComet Software [17] %TP	%FN	%FP	PPV	Sensitivity
Case 1	60	43	18.6	81.4	558.14	0.032	18.6
Case 2	128	104	92.31	7.69	98.08	0.485	92.31
Case 3	85	76	44.74	55.26	223.68	0.167	44.74
Case 4	50	38	78.95	21.05	157.89	0.333	78.95
Case 5	106	74	85.14	14.86	151.35	0.36	74
Average	429	335	**63.95**	**36.05**	**237.83**	**0.275**	**63.95**

Table 3. Performance analysis of the Proposed Method

Cases	Total No. of Comets	TNC	The Proposed Method %TP	%FN	%FP	PPV	Sensitivity
Case 1	60	43	81.4	18.6	16.28	0.833	81.4
Case 2	128	104	95.19	4.81	0.962	0.99	95.19
Case 3	85	76	97.37	2.63	2.63	0.974	97.37
Case 4	50	38	100	0	0	1	100
Case 5	106	74	91.89	8.11	4.05	0.958	91.89
Average	429	335	**93.17**	**6.83**	**4.784**	**0.951**	**93.17**

Table 4. Performance analysis of the proposed method compared with other two methods

Parameters	Sansone's *et al.* method [14]	OpenComet Software [17]	The Proposed Method
% TP	78.96	63.95	93.17
% FN	21.04	36.05	6.83
% FP	14.03	237.83	4.78
PPV	0.871	0.275	0.951
Sensitivity	78.96	63.95	93.17

5. Conclusion & future work

An improved method for automatic detection of comets in silver stained comet assay images for DNA damage analysis is developed and implemented in this paper. With the proposed method, we were able to select the region of interest with minimum loss of information in the comet tail regions. The proposed method gives better results in terms of %TP, %FN, %FP, PPV and sensitivity when compared with the most recent and related method available in the literature. In future, a better method for eliminating overlapping nuclei has to be developed.

References

1. Ostling O, Johanson K. Microelectrophoretic study of radiation-induced dna damages in individual mammalian cells. *Biochem Biophys Res Commun* 1984; **123**(1):291-298.
2. Singh NP, McCoy MT, Tice RR, Schneider EL. A simple technique for quantitation of low levels of dna damage in individual cells. *Exp Cell Res* 1988; **175**(1):184-191.
3. Dhawan A, Mathur N, Seth PK. The effect of smoking and eating habits on dna damage in indian population as measured in the comet assay. *Mutat Res/Fundam Mol Mech Mutagen* 2001; **474**(1):121-128.
4. Jackson SP, Bartek J. The dna-damage response in human biology and disease. *Nature* 2009; **461**(7267):1071-1078.
5. Olive PL, Banáth JP. The comet assay: a method to measure dna damage in individual cells. *Nat Protoc* 2006; **1**(1):23-29.
6. Tice R, Agurell E, Anderson D, Burlinson B, Hartmann A, Kobayashi H. Single cell gel/comet assay: guidelines for in vitro and in vivo genetic toxicology testing. *Environ Mol Mutage* 2000; **35**(3):206-221.
7. Nossoni F. Single-cell gel electrophoresis (comet assay): methodology, potential applications, and limitations in cancer research. *MMG 445 Basic Biotechnol EJ* 2008; **4**(1):30-35.
8. Nandhakumar S, Parasuraman S, Shanmugam M, Rao KR, Chand P, Bhat BV. Evaluation of dna damage using single-cell gel electrophoresis (comet assay). *J Pharmacol Pharmacother* 2011; **2**(2):107-111.
9. Nadin SB, Vargas-Roig LM, Ciocca DR. A silver staining method for single-cell gel assay. *J Histochem Cytochem* 2001; **49**(9):1183-1186.
10. Rivest JF, Tang M, McLean J, Johnson F. Automated measurements of tails in the single cell gel electrophoresis assay. *Quality Measurements: The Indispensable Bridge between Theory and Reality* 1996; **1**:111-114.
11. Helma C, Uhl M. A public domain image-analysis program for the single-cell gel-electrophoresis (comet) assay. *Mutat Res/Genet Toxicol Environ Mutagen* 2000; **466**(1):9-15.
12. Końca K, Lankoff A, Banasik A, Lisowska H, Kuszewski T, Góźdź S. A cross-platform public domain pc image-analysis program for the comet assay. *Mutat Res/Genet Toxicol Environ Mutagen* 2003; **534**(1):15-20.
13. González J, Romero I, Barquinero J, García O. Automatic analysis of silver-stained comets by cellprofiler software. *Mutat Res/Genet Toxicol Environ Mutagen* 2012; **748**(1):60-64.
14. Sansone M, Zeni O, Esposito G. Automated segmentation of comet assay images using gaussian filtering and fuzzy clustering. *Med Biol Eng Comput* 2012; **50**(5):523-532.
15. Böcker W, Rolf W, Bauch T, Müller WU, Streffer C. Automated comet assay analysis. *Cytometry* 1999; **35**(2):134-144.
16. Frieauff W, Hartmann A, Suter W. Automatic analysis of slides processed in the comet assay. *Mutagenesis* 2001; **16**(2):133-137.
17. Gyori BM, Venkatachalam G, Thiagarajan P, Hsu D, Clement MV. Opencomet: An automated tool for comet assay image analysis.*Redox Biol.* 2014; **2**:457-465.
18. Gonzalez RC, Richard EW. *Digital image processing*. 3rd ed. India: Pearson Education; 2008.
19. Otsu N. A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern* 1979; **9**(1):62-66.