International Journal of Pattern Recognition
and Artificial Intelligence
Vol. 30, No. 5 (2016) 1657002 (16 pages)

DOI: 10.1142/S0218001416570020

World Scientific
www.worldscientific.com

A Simple and Efficient Arrowhead Detection Technique in Biomedical Images

K.C. Santosh*

Department of Computer Science, The University of South Dakota
414 E Clark St, Vermillion, SD 57069, USA
santosh.kc@usd.edu

Naved Alam† and Partha Pratim Roy‡

Department of Computer Science, Indian Institute of Technology
Roorkee, Uttarakhand 247667, India
†*navedalamalam@gmail.com*
‡*proy.fcs@iitr.ac.in*

Laurent Wendling

LIPADE, Université Paris Descartes (Paris V)
75270 Paris Cedex 06, France
laurent.wendling@parisdescartes.fr

Sameer Antani§ and George R. Thoma¶

National Library of Medicine, National Institutes of Health
8600 Rockville Pike, Bethesda, MD 20894, USA
§*sameer.antani*
¶*george.thoma@nih.gov*

Received 11 November 2015
Accepted 5 January 2016
Published 1 April 2016

In biomedical documents/publications, medical images tend to be complex by nature and often contain several regions that are annotated using arrows. In this context, an automated arrowhead detection is a critical precursor to region-of-interest (ROI) labeling and image content analysis. To detect arrowheads, in this paper, images are first binarized using fuzzy binarization technique to segment a set of candidates based on connected component (CC) principle. To select arrow candidates, we use convexity defect-based filtering, which is followed by template matching via dynamic time warping (DTW). The DTW similarity score confirms the presence of arrows in the image. Our test results on biomedical images from imageCLEF 2010 collection shows the interest of the technique, and can be compared with previously reported state-of-the-art results.

Keywords: Arrowhead detection; document images; biomedical publications; image region labeling; content-based image retrieval.

* Corresponding author.

1. Introduction

1.1. *Motivation*

Essential information is often conveyed succinctly through graphical illustrations and figures/images in biomedical publications. Medical images tend to be complex by nature, and are often annotated with graphical overlay pointers, such as arrows and asterisk. Medical researchers often use these pointers to highlight meaningful regions-of-interest (ROIs) (see Figs. 1 and 2), while minimizing distractions from other less relevant regions. Additionally, they are often referred to figure captions and mentioned in the paper. Therefore, detecting arrows could help identify meaningful ROIs and annotate them with the concepts appearing in the biomedical text.[5,6] In Fig. 1, we provide a complete scenario of the project where the importance of the arrow is highlighted. This paper improves on prior work in arrow detection toward meeting this goal in image content analysis.

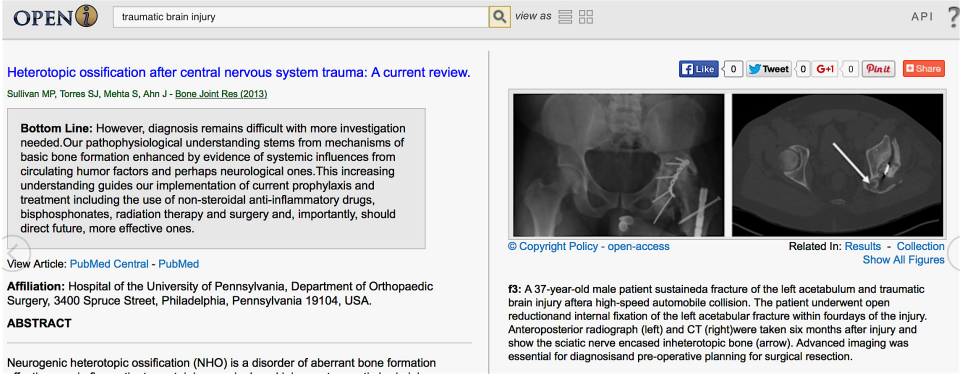
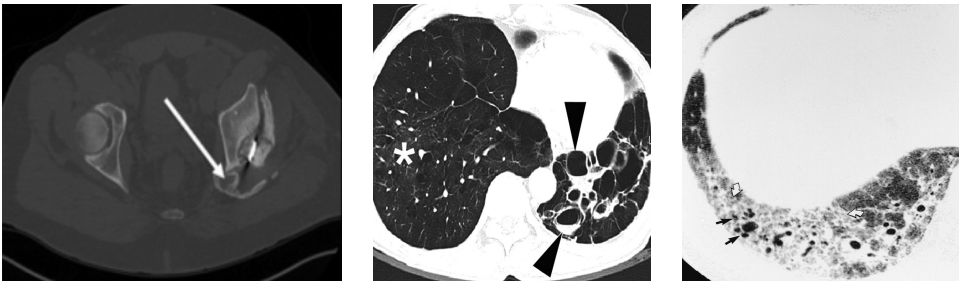


Fig. 1. Using NLM's open-i image retrieval search engine (url: https://openi.nlm.nih.gov), the illustration highlights the importance of using arrow in biomedical images (i.e. its location pointing ROI and relationship between the texts and ROI).



(a) (b) (c)

Fig. 2. Three examples showing different types of arrows pointing specific image regions. These are taken from published biomedical papers (see Fig. 1). Arrow types can be just a triangle (i.e. a regular arrowhead) or with straight and curved tail. In addition, their intensities vary with respect to the background.

1.2. *Related work*

We find that there are few techniques reported in the literature to detect overlaid arrows. Existing methods rely on sparse pixel vectorization, segmenting text-like and symbol-like objects, and global or local thresholding.

Dori and Wenyin[8] proposed a technique to detect arrows based on sparse pixel vectorization.[7] The concept relies on the cross-sectional runs (or width runs) of black image regions (assuming arrow in black). These runs represent the line at intervals along the tracking direction and records the middle points of these sections. The points are then used to construct vectors. The vectorization process results in many thick short bars from the arrow heads that are then used to make a decision. The technique utilizes an interesting application but is limited to machine printed line images. Features such as eccentricity, convex area and solidity has been used to detect arrows, but the current techniques are limited to regular arrows (i.e. straight arrows showing left, right, top and bottom).[16] Additionally, the method uses predefined threshold to avoid small objects and noise. Cheng *et al.* use text-like and arrow-like objects separation, assuming that arrows are shown in either black or white color with respect to the background.[2] From the binary image, arrow-like object separation employs a fixed sized mask (after removing the small objects and noise as in Ref. 16), which are then used for feature computation such as major and minor axis lengths, axis ratio, area, solidity and Euler number. A recent study uses a pointer region and boundary detection to handle distorted arrows,[25] which is followed by edge detection techniques and fixed thresholds as reported in Refs. 23 and 24. These candidates are used to compute overlapping regions, which are then binarized to extract the boundary of the expected pointers.

Fundamentally, edge-based arrow detection techniques are limited by the weak-edge problem.[2,16,25] No matter how robust the arrow detection techniques are, hard thresholding (either global or local) is one of the primary reasons for failure. This means that a hard threshold cue often weakens the decision in pointer detection. For edge detection in binary or grayscale images, most state-of-the-art methods use classical algorithms like Roberts, Sobel and Canny edge detection. Template-based methods are limited since they require new templates to train new images. Also, it may be necessary to re-evaluate the threshold values when new images are used. Edge-based techniques are still considered since sampling points can be remarkably compact compared to solid regions, especially when broken boundaries are recoverable. In biomedical images, one of the major issues for a broken boundary is the nonhomogeneous intensity distribution, where pointers overlap with content. According to Hori and Doermann,[11] broken lines can be recovered when gaps are small but, in practice, they are often inaccurate.[11]

1.3. *Contribution outline*

Our method (OM) can be summarized as shown in Fig. 3. It relies on a grayscale fuzzy binarization process at different levels, since straight forward thresholding may

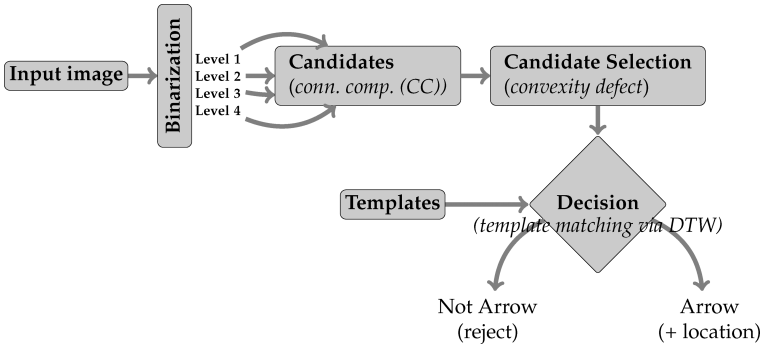


Fig. 3. Overall system workflow in block format. Block-wise explanation can be found in Sec. 2.

not work (see Sec. 3). Similar to previously reported work,[22] candidates are segmented based on connected component (CC) principle. These candidates are filtered using hull convexity defect-based technique. This step helps prune artifacts (or unwanted noisy CCs) and store arrowhead-like candidates. Next we perform template matching using dynamic programming (i.e. dynamic time warping (DTW)) to confirm whether the candidate is an arrowhead. In our assessment, an arrow is said to be detected if their matching score exceeds an empirically set threshold.

In this paper, unlike the common state-of-the-art methods, OM uses four different levels of fuzzy binarization. This ensures that overlaid arrow candidates are not missed. However, it may result in repetitions. We note that the primary variation in an arrow appearance is due to its tail (shape and size). Therefore, OM limits itself to just detecting arrowheads, which is the extension of the proceedings presented in the imaging symposium 2016.[20]

The remainder of the paper is organized as follows. In Sec. 2, we explain OM in detail, where it mainly includes binarization process (Sec. 2.1) and candidate selection (Sec. 2.2). Results are reported in Sec. 3, including a comprehensive state-of-the-art comparison. In Sec. 4, we state conclusions and provide next-steps.

2. Method

2.1. *Binarization*

In biomedical images (see Fig. 2), arrows appear with either high or low intensity to enhance their visibility in the image. In addition, in many cases arrows are blurred, overlapped or surrounded by textured areas. In such contexts, typical binarization tools that are based on fixed threshold values are unable to perfectly extract candidate regions. Therefore, we focus on an adaptive binarization tool, which is based on a fuzzy partition of a two-dimensional (2D) histogram of the image, taking into account the gray level intensities and local variations.[4,3] Two-dimensional Z-function criteria based on the optimization of fuzzy entropy are then computed from

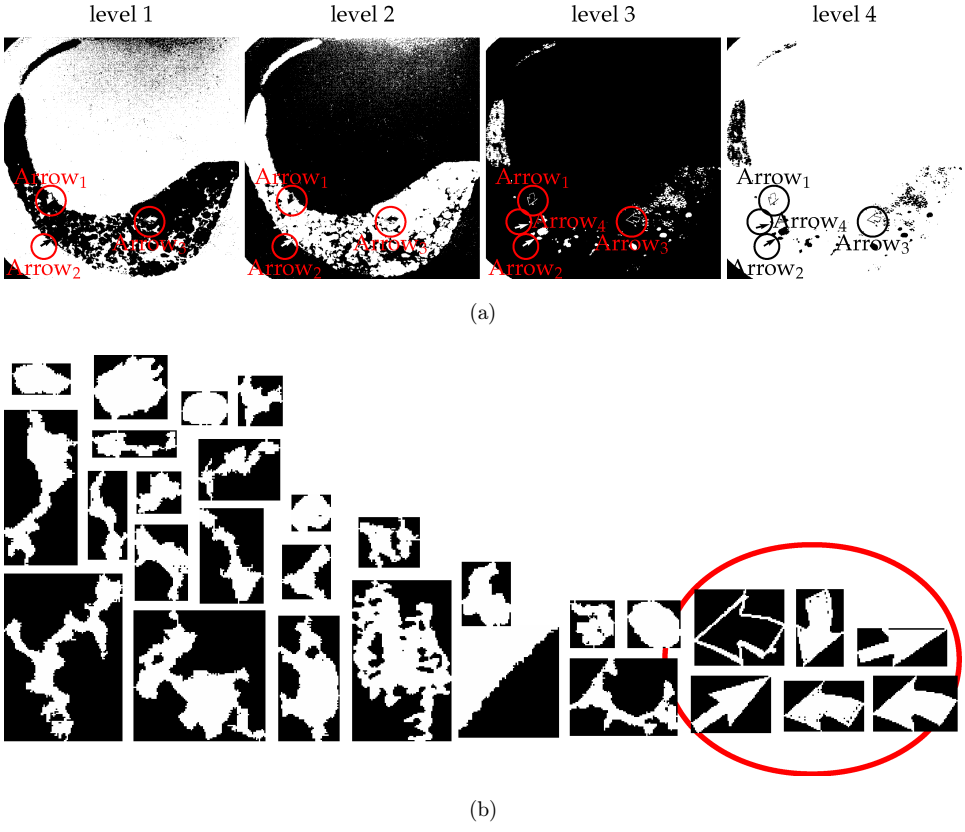


Fig. 4. (Color online) Fuzzy binarization (of Fig. 2(c)): (a) four different levels (level 1 to level 4), where the segmented arrows are encircled both in red and black with respect to the background color; and (b) a collection of all segmented CCs including arrows (encircled in red).

this histogram to automatically set the threshold.[3] The Z-function employs two kernels: low level and high level cuts, in addition to direct inversions. The latter issue (image inversion) takes opposite image intensities into account. Altogether, four different binarized levels are processed, as illustrated in Fig. 4. In Fig. 4(a), arrow candidates are encircled in both red and black (with respect to the background color). The main idea of using four different levels of binarization is not to miss the overlaid arrows. Furthermore, deformed and/or distorted arrows can be discarded since the arrows are repeated in other levels of binarization. In Fig. 4(b), some of the arrows are repeatedly segmented. In contrast, straight forward thresholding technique (Otsu,[15] for instance) may not be able to segment those arrows because of the large intensity variations from one image to another.

From a pool of several candidates (see Fig. 4(b)), we are required to select arrow-like candidates. In what follows, we describe a complete candidate selection process in detail.

2.2. *Candidate selection*

Our candidate selection process is based on the characteristics of the arrowhead, which can typically be represented by a triangle. Unlike the previously reported work,[22] we do not take tail information into account. One of the primary reasons is that it may vary geometric signatures computed from extreme points of a triangle (i.e. triplet) because tail structures tend to vary from time to time. Such a change will affect overall appearance of the arrow (Fig. 5). After we detect arrowhead, we will take the corresponding tail into account since both came from the same CC.

To detect an arrowhead, the following steps are carried out:

(1) convexity defect-based arrowhead candidate cropping; and
(2) arrowhead candidate matching with the templates.

2.2.1. *Convexity defect-based arrowhead candidate cropping*

To select arrow-like candidates, we apply hull convexity defect concept (see Fig. 6). A set of points along the contour of the binary CC are defined to be convex if it contains the line segments connecting each pair of its points. In a convex combination, each point x_i in the set S is assigned a weight or coefficient w_i in such a way that the coefficients are all non-negative and sum to one, and these weights are used to

Fig. 5. Examples showing the changes in tail structure. Further, an absence of the tail is also possible.

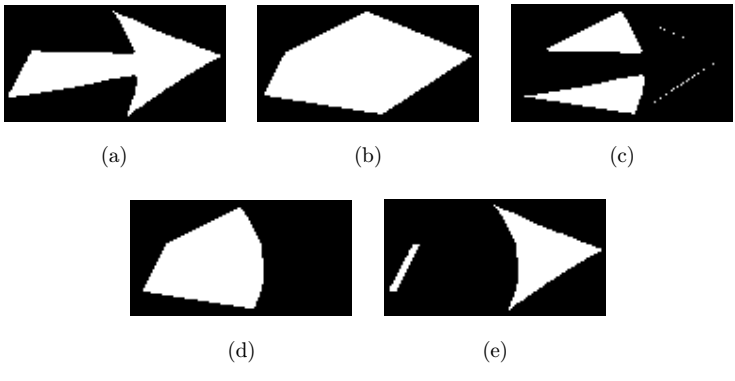


Fig. 6. Arrowhead candidate cropping: (a) An arrow, (b) convex hull, (c) convexity defect, (d) a complete convexity defect region and (e) arrowhead candidates.

compute a weighted average of the points. For each choice of coefficients, the resulting convex combination is a point in the convex hull, and the whole convex hull can be formed by choosing coefficients in all possible ways. Expressing this as a single formula, the convex hull is the set:

$$\left\{ \sum_{i=1}^{|S|} w_i x_i | (\forall i : w_i \geq 0) \wedge \sum_{i=1}^{|S|} w_i = 1 \right\}. \tag{1}$$

This means that the convex hull of a finite point set $S \in \mathbb{R}^n$ forms a convex polygon when $n = 2$. In Fig. 6(b), an example is shown. Using such a convex hull, we attempt to remove the tail since their exists convex shaped silhouettes in both sides (see Fig. 6 (c)), which is computed by subtracting an original candidate from the convex hull. In Fig. 6(d), the convexity defect region is shown, which is just a convex hull of both convex shaped silhouettes. At the end, in Fig. 6(e), arrowhead candidate(s) is(are) selected by subtracting an original image with the convexity defect region.

2.2.2. *Arrowhead candidate matching with template*

To confirm arrowhead candidates (see Fig. 6), we apply a template matching technique. We extract a feature along the contour and match with the predefined templates using DTW technique. The arrowhead candidate is confirmed when the similarity score crosses the empirically designed threshold.

Feature extraction. Along the contour, we have a set of coordinate points,

$$P = \{p_i\}_{i=1,\ldots,n}, \tag{2}$$

where $p_i = (x_i, y_i)$. To extract feature vector (f), we compute the change in angle with respect to x-axis from any consecutive pair,

$$f = \{\alpha_i\}_{i=1,\ldots,n}, \tag{3}$$

where $\alpha_i = \arctan\left(\frac{y_i - y_{i-1}}{x_i - x_{i-1}}\right)$. This goes in a cyclic order either clockwise or anti-clockwise but, following the trigonometry, we follow anti-clockwise traversal. In our feature vector, continuous redundancy of α_i can be possible, $\alpha_i = \alpha_{i+j}, j = 1, \ldots, m$, where $m \leq n$. Therefore, it is desired to express the contours of shapes with a few representative pixels (called the dominant points). Through polygonal approximation,[18,9,17] we represent a digital curve using fewer points such that the properties of the curvature of the digital curve are retained. Next the geometrical properties like inflexion points or concavities can be evaluated. Besides, to make it simple and effective, we compute the difference between the angles and check whether it crosses the threshold, ϵ. The choice of ϵ is usually user-defined. This means we take α_i if $|\alpha_i - \alpha_{i+1}| \leq \epsilon$ (in our case, $\epsilon = 0$). Like most line fitting/polygonal approximation (or dominant point detection) methods, it can be made nonparametric by using the error bound due to digitization as a termination condition. Figure 7 shows three examples, where the changes in angles are shown at all dominant points. To make the

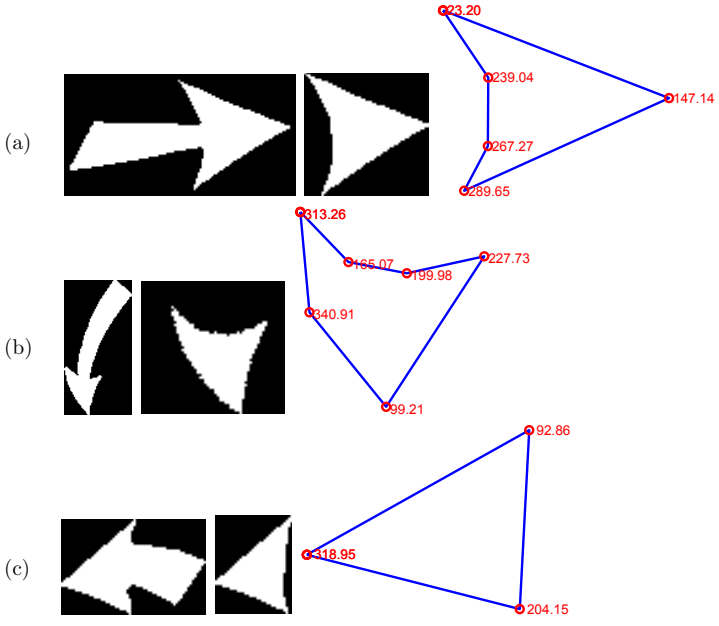


Fig. 7. Three examples showing a complete process (from left to right) starting from an original candidate (resulting from fuzzy binarization — see Fig. 4), arrowhead cropping (see Fig. 6) to feature extraction after polygonal approximation.

feature vector rotation invariant, one needs to follow either clockwise or anti-clockwise to compute changes in angles.

Dynamic time warping. DTW allows to find the dissimilarity between two non-linear sequences potentially having different lengths.[12,19] In Fig. 7, one can notice the variations in feature vector from one arrowhead to another. Let us consider two feature sequences:

$$f_1 = \{\alpha_i\}_{1=1,\ldots,n} \quad \text{and} \quad f_2 = \{\beta_j\}_{j=1,\ldots,m}, \tag{4}$$

of size n and m, respectively. The aim of the algorithm is to provide the optimal alignment between both sequences. At first, a matrix of size $n \times m$ is constructed. Then for each element, local distance metric $\delta(i,j)$ between the events e_i and e_j is computed, i.e. $\delta(i,j) = (e_i - e_j)^2$. Let $D(i,j)$ be the global distance up to (i,j),

$$D(i,j) = \min \begin{bmatrix} D(i-1,j-1), \\ D(i-1,j), \\ D(i,j-1) \end{bmatrix} + \delta(i,j) \tag{5}$$

with an initial condition $D(1,1) = \delta(1,1)$ such that it allows warping path going diagonally from starting node $(1,1)$ to end (n,m). The main aim is to find the path for which the least cost is associated. The warping path therefore provides the difference cost between the compared signatures. Formally, the warping path is,

$W = \{w_k\}_{k=1,\ldots,l}$, where $\max(i,j) \leq l < i + j - 1$ and kth element of W is $w(i,j)_k \in [1:n] \times [1:m]$ for $k \in [1:l]$. The optimized warping path W satisfies the following three conditions:

- boundary condition: $w_1 = (1,1)$ and $w_l = [n,m]$,
- monotonicity condition: $i_1 \leq i_2 \leq \cdots \leq i_n$ and $j_1 \leq j_2 \leq \cdots \leq j_m$, and
- continuity condition: $w_{k+1} - w_k \in \{(1,1)(0,1)(1,0)\}$ for $k \in [1:l-1]$.

Boundary condition conveys that the path starts from $(1,1)$ to (n,m), aligning all elements to each other. Monotonicity condition forces the path advances one step at a time. Continuity condition restricts allowable steps in the warping path to adjacent cells, and therefore it does not go backward. Note that continuity condition implies monotonicity condition. We then define the global distance between f_1 and f_2 as,

$$\Delta(f_1, f_2) = \frac{D(n,m)}{l}. \tag{6}$$

The last element of the $n \times m$ matrix gives the DTW-distance between f_1 and f_2, which is normalized by l, i.e. the number of discrete warping steps along the diagonal DTW-matrix. Overall, DTW measures the similarity between two sequences, and can be summarized as follows.

(1) Thanks to DTW, noise in arrowhead (along the contour) does not let the cost to go beyond the threshold. This means that some of the arrowheads with noisy artifacts connected to them are still detected.
(2) If the cropped candidate is not actually an arrowhead, DTW results in high cost.

Further, feature extraction and DTW matching techniques provide robustness to the change in rotation. As an example, $\Delta($  ,  $) = 0.00$ and $\Delta($,  $) = 0.00.$, where in both cases the second arrow is rotated by 30° and 120°, respectively. It holds the same for image scaling (i.e. robust to scaling).

3. Experiments

3.1. *Datasets, ground-truth and evaluation protocol*

The well-known imageCLEF dataset[14] is used for testing. It is composed of 298 chest CT images. Each image is expected to have at least one arrow, and there are 1049 pointers, in total. For all images in the dataset, ground-truths of the pointers were created and each ground-truth includes information like arrow type, color, location and direction. For validation, for any given image in the dataset, our performance evaluation criteria are precision, recall and F_1 score,

$$\text{precision} = \frac{m_1}{M}, \quad \text{recall} = \frac{m_1}{N} \quad \text{and}$$

$$F_1 \text{ score} = 2\left(\frac{(m_1/M) \times (m_1/N)}{(m_1/M) + (m_1/N)}\right), \tag{7}$$

where m_1 is the number of correct matches from the detected set M and N is the total number of pointers (in the ground-truth) that are expected to be detected. Precision defines whether the retrieved candidate is relevant (i.e. an arrow), and recall defines how relevant is the search.

3.2. *Our result and analysis*

Table 1 shows the performance evaluation scores in terms of precision, recall and F_1 score. In the reported results, we prioritize the recall measure since we do not like to miss arrow candidates. The method achieved F_1 score of 91.09%. As mentioned in Sec. 1.3, we also reported the results using Otsu thresholding[15] in place of fuzzy binarization. The results shown in Table 1 confirm the usefulness of fuzzy binarization at four different levels. Considering fuzzy binarization, it is important to note that while checking with the ground-truths, we are able to segment 95.1% (recall value) of arrows from our complete dataset. In contrast, we have a recall value of 84.81% from Otsu. This means that we have missed more than 15% of the arrows during segmentation, which is not the case for fuzzy binarization.

Our method (OM1 in Table 1) is able to detect arrowheads regardless of their tail structure. But, if the shape of the arrowhead is affected by noisy artifacts, the proposed method fails. Figure 8 shows both examples: noisy artifacts that are connected along the tail, and noisy artifacts that are connected with arrowhead. Also, the method does not detect highly curved arrows since convexity defect-based arrowhead cropping does not yield expected arrowhead candidates.

3.3. *Comparative study*

Further, the comparative study with state-of-the-art methods has been made. In this comparison, our benchmarking methods are categorized into two groups:

(1) state-of-the-art methods that are specially designed for arrow detection; and
(2) common template-based method by using well-known state-of-the-art shape descriptors.

Table 1. Performance (in %) of the proposed method.

Our Method	Precision	Recall	F_1 Score
OM1	88.50	93.80	91.09
OM2	68.21	74.32	71.13

Index
OM1: out method using fuzzy binarization[3] (four different image levels).
OM2: our method using Otsu[15] (single image level).

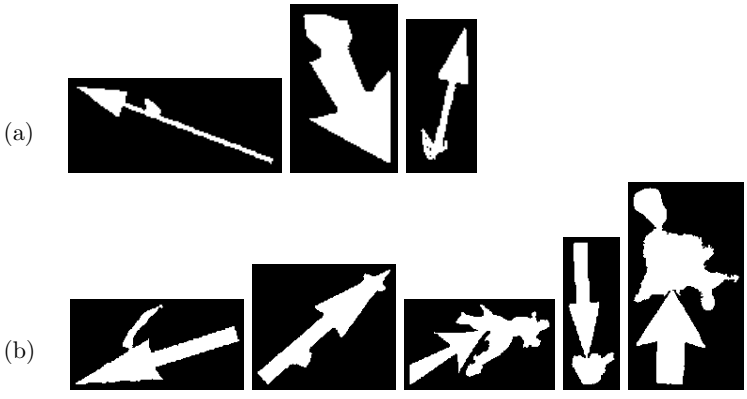


Fig. 8. Examples of when the proposed method succeeds and fails: (a) Noisy artifacts connected along the tail do not affect the method and (b) they do largely affect when connected with arrowhead.

3.3.1. *Arrow detection methods*

Four well-known methods from the state-of-the-art that are specially designed for arrow detection are used:

(1) global thresholding-based method,[2]
(2) two edge-based methods,[24,25] and
(3) a template-free geometric signature-based method.[22]

The results are provided in Table 2, where Santosh *et al.*[22] performs the best with precision, recall and F_1 score values 93.14%, 86.93% and 89.94%, respectively. In our comparison, to avoid biasing, these performance scores are taken from their published articles.

3.3.2. *Template-based methods*

In case of template-based method, we created 11 templates (arrows) having different shapes (including sizes). The template size can further be extended in accordance with the dataset. To extract shape features, we took the most frequently used shape descriptors (in computer vision) from the state-of-the-art. They are

(1) generic Fourier descriptor (GFD),[26]
(2) shape context (SC),[1]

Table 2. Performance (in %) of the previously reported methods.

Previously Reported Methods	Precision	Recall	F_1 Score
Cheng *et al.*[2]	81.10	74.10	77.00
You *et al.*[24]	22.80	77.80	35.00
You *et al.*[25]	84.20	81.60	83.00
Santosh *et al.*[22]	93.14	86.92	89.94

Table 3. Performance (in %) of the template-based methods.

Template-Based Methods	Precision	Recall	F_1 Score
GFD[26]	75.10	78.33	76.68
SC[1]	68.30	71.40	69.82
ZM[13]	55.20	57.70	56.40
RT[10]	59.50	63.60	61.48
D-Radon[21]	62.10	65.30	63.65

(3) Zernike moment (ZM),[13]
(4) R-transform (RT)[10] and
(5) DTW-Radon.[21]

For these descriptors, it is important to fit the best parameters. In case of GFD, one needs to tune the radial (4:12) and angular (6:20) frequency parameters to get the best combinations. Note that such a best combination of radial and angular frequencies can be different from one dataset to another. For SC, we used the 100 sample points as reported in Ref. 1 by omitting smaller CCs. In case of ZM, we applied 36 Zernike functions of order less than or equal to 7. For the Radon transform, projecting range is $[0, \pi)$. These shape descriptors are rotation-, scale- and translation-invariant, and thus useful in our application since CCs are observed at different sizes, scales and directions. After extracting features, the main idea is to rank the CCs from any studied image based on the order of shape similarity. In our test, top-10 ranking has been implemented. Using this framework, results (precision, recall and F_1 score) are provided in Table 3. Among all shape descriptors, GFD provides the best performance.

3.3.3. *Best of the worlds comparison*

In Fig. 9, we compare the best scores from three different studies/results reported in Tables 1–3. On the whole, considering such a dataset, the proposed method outperforms the best state-of-the-start arrow detection method by more than 1% F_1

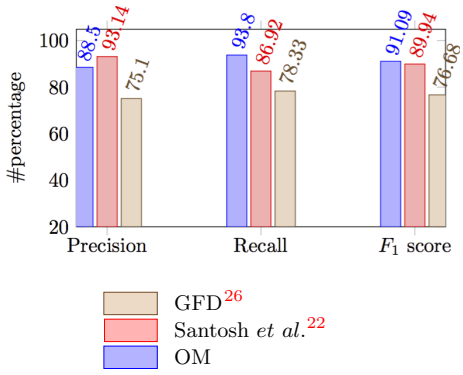


Fig. 9. Performance comparison. The highest scores are taken from all studies (see Tables 1–3).

score, and the template-based (shape descriptor) method by more than 16% F_1 score, at the cost of low precision and high recall.

4. Conclusion and Future Work

In this paper, we have presented a new method to detect overlaid arrows in biomedical images. Images are first binarized via fuzzy binarization tool to segment a set of candidates. To select arrow candidates, we use a hull convexity defect-based arrowhead cropping, which is followed by template matching via dynamic programming. In our assessment, (using imageCLEF 2010 collection), our results outperforms the state-of-the-art methods.

To the best of our knowledge, this is the first time arrow detection has been done without using tail information since variations in the shape and size of the tail change an overall shape of the complete arrow. As our next steps, we plan to integrate previously reported techniques (state-of-the-art methods) that can be used as pre- or post-processing steps. Further, use of machine learning instead of using template-based approaches, would be our immediate concern. Also, since the current work does not use any color information (that could possible be appeared in biomedical images), we plan to adopt them in our study.

Acknowledgments

This research was supported [in part] by the Intramural Research Program of the National Institutes of Health (NIH), National Library of Medicine (NLM) and Lister Hill National Center for Biomedical Communications (LHNCBC).

References

1. S. Belongie, J. Malik and J. Puzicha, Shape matching and object recognition using shape contexts, *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(4) (2002) 509–522.
2. B. Cheng, R. J. Stanley, S. De, S. Antani and G. R. Thoma, Automatic detection of arrow annotation overlays in biomedical images, *Int. J. Healthc. Inf. Syst. Inform.* **6** (2011) 23–41.
3. H. Cheng and Y.-H. Chen, Fuzzy partition of two dimensional histogram and its application to thresholding, *Pattern Recognit.* **32** (1999) 825–843.
4. Z. Chi, H. Yan and T. Pham, *Fuzzy Algorithms: With Applications to Image Processing and Pattern Recognition* (World Scientific Publishing River Edge, NJ, USA, 1996).
5. D. Demner-Fushman, S. Antani, M. Simpson and M. Rahman, Combining text and visual features for biomedical information retrieval and ontologies, Technical Report LHNCBC Board of Scientific Counselors, National Institutes of Health (Bethesda, MD, 2010).
6. D. Demner-Fushman, S. Antani, M. S. Simpson and G. R. Thoma, Design and development of a multimodal biomedical information retrieval system, *J. Comput. Sci. Eng.* **6**(2) (2012) 168–177.
7. D. Dori, S. Member and W. Liu, Sparse pixel vectorization: An algorithm and its performance evaluation, *IEEE Trans. Pattern Anal. Mach. Intell.* **21** (1999) 202–215.
8. D. Dori and L. Wenyin, Automated cad conversion with the machine drawing understanding system: Concepts, algorithms, and performance, *IEEE Trans. Syst. Man Cybern. A, Syst. Humans* **29** (1999) 411–416.

9. D. H. Douglas and T. K. Peucker, Algorithms for the reduction of the number of points required to represent a digitized line or its caricature, *Can. Cartograph.* **10**(2) (1973) 112–122.
10. T. V. Hoang and S. Tabbone, The generalization of the r-transform for invariant pattern representation, *Pattern Recognit.* **45**(6) (2012) 2145–2163.
11. O. Hori and D. S. Doermann, Robust table-form structure analysis based on box-driven reasoning, in *Document Analysis and Recognition, 1995., Proceedings of the Third International Conference on*, Vol. 1 (1995), pp. 218–221.
12. E. J. Keogh and M. J. Pazzani, Scaling up dynamic time warping to massive dataset, in *Proceedings of the Third European Conference on Principles of Data Mining and Knowledge Discovery, PKDD'99* (Springer-Verlag, London, UK, 1999), pp. 1–11.
13. W.-Y. Kim and Y.-S. Kim, A region-based shape descriptor using zernike moments, *Signal Process., Image Commun.* **16**(1–2) (2000) 95–102.
14. H. Müller, J. Kalpathy-Cramer, I. Eggel, S. Bedrick, S. Radhouani, B. Bakke, C. E. Kahn Jr and W. Hersh, Overview of the clef 2009 medical image retrieval track, in *Multilingual Information Access Evaluation II. Multimedia Experiments* (Springer, New York, 2010), pp. 72–84.
15. N. Otsu, A Threshold selection method from gray-level histograms, *IEEE Trans. Syst. Man, Cybern.* **9**(1) (1979) 62–66.
16. J. Park, W. Rasheed and J. Beak, Robot navigation using camera by identifying arrow signs, in *Int. Conf. Grid and Pervasive Computing — Workshops* (IEEE Computer Society, 2008), pp. 382–386.
17. D. K. Prasad, M. K. Leung, C. Quek and S.-Y. Cho, A novel framework for making dominant point detection methods non-parametric, *Image Vis. Comput.* **30**(11) (2012) 843–859.
18. U. Ramer, An iterative procedure for the polygonal approximation of plane curves, *Comput. Graph. Image Process.* **1**(3) (1972) 244–256.
19. H. Sakoe, Dynamic programming algorithm optimization for spoken word recognition, *IEEE Trans. Acoust. Speech Signal Process.* **26** (1978) 43–49.
20. K. C. Santosh, N. Alam, P. P. Roy, L. Wendling, S. Antani and G. R. Thoma, Arrowhead detection in biomedical images, in *Document Recognition and Retrieval (XXIII), IS&T Electronic Imaging Symp.* (2016) to appear.
21. K. C. Santosh, B. Lamiroy and L. Wendling, Dtw-radon-based shape descriptor for pattern recognition, *Int. J. Pattern Recognit. Artif. Intell.* **27**(3) (2013) 1350008.
22. K. C. Santosh, L. Wendling, S. Antani and G. Thoma, Scalable arrow detection in biomedical images, in *Proc. IAPR Int. Conf. Pattern Recognition* (IEEE Computer Society, Stockholm (Sweden), August 2014), pp. 3257–3262.
23. D. You, S. Antani, D. Demner-Fushman, M. M. Rahman, V. Govindaraju and G. R. Thoma, Biomedical article retrieval using multimodal features and image annotations in region-based cbir, in *Document Recognition and Retrieval*, eds. L. Likforman-Sulem and G. Agam, *SPIE Proc.* Vol. 7534 (SPIE, 2010), pp. 1–10.
24. D. You, E. Apostolova, S. Antani, D. Demner-Fushman and G. R. Thoma, Figure content analysis for improved biomedical article retrieval, in *Document Recognition and Retrieval*, eds. K. Berkner and L. Likforman-Sulem, *SPIE Proc.* Vol. 7247 (SPIE, 2009), pp. 1–10.
25. D. You, M. S. Simpson, S. Antani, D. Demner-Fushman and G. R. Thoma, A robust pointer segmentation in biomedical images toward building a visual ontology for biomedical article retrieval, in *Document Recognition and Retrieval*, eds. R. Zanibbi and B. Coüasnon, *SPIE Proc.* Vol. 8658 (SPIE, 2013).
26. D. Zhang and G. Lu, Shape-based image retrieval using generic fourier descriptor, *Signal Process., Image Commun.* **17** (2002) 825–848.

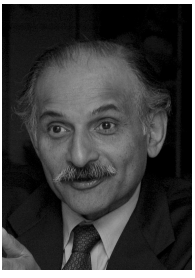
K. C. Santosh, Ph.D., is an Assistant Professor at the Department of Computer Science at the University of South Dakota (USD). Before joining the USD, from 2013 to 2015, he worked as a Research fellow at the U.S. National Library of Medicine (NLM), National Institutes of Health (NIH). He worked as a Postdoctoral research scientist at the LORIA research centre, Universite de Lorraine in direct collaboration with industrial partner ITESOFT, France, for two years. He also worked as a Research scientist at the INRIA Nancy Grand Est research centre for three years, until 2011. He has demonstrated expertise in pattern recognition, image processing, computer vision and machine learning with various applications in handwriting recognition, graphics recognition, document information content exploitation, medical image analysis and biometrics.

Naved Alam is currently a Master student in Computer Science at Indian Institute of Technology (IIT), Roorkee. His research areas include image analysis, pattern recognition and machine learning.

Partha Pratim Roy received his Ph.D. in Computer Science in 2010 from Universitat Autònoma de Barcelona, Barcelona. He worked as postdoctoral research fellow in the Computer Science Laboratory (LI, RFAI group), France (2010–2012) and in Synchromedia Lab, Canada (2013). Presently, Dr. Roy is working as an Assistant Professor at Indian Institute of Technology (IIT), Roorkee. His main research area is pattern recognition.

Laurent Wendling received his Ph.D. in Computer Science in 1997 from the Université Paul Sabatier, Toulouse, France. He received his Habilitation degree in 2006. From 1999 to 2009, he was an Assistant Professor at the ESIALNancy, and a member of LORIA. He is currently a Full Professor at the Université Paris Descartes. He is also a member of the SIP team, LIPADE. His current research topics are spatial relations, feature selection and image segmentation.

Sameer Antani is a Staff Scientist at the National Library of Medicine (NLM), part of the National Institutes of Health (NIH), USA. His research applies his expertise in biomedical imaging and analysis, informatics, computer vision, machine learning, computer science and engineering technology to advance the biomedical sciences, global health and the dissemination of scientific information. He is a Senior member of the SPIE, member of IEEE, and serves as Vice Chair for Computational Medicine in IEEE Computer Society's Technical Committee on Computational Life Sciences. Dr. Antani has a Ph.D. and a M. Eng. degree in Computer Science and Engineering from the Pennsylvania State University, USA, and B. Eng. degree in Computer Engineering from the University of Pune, India. See: http://go.usa.gov/4wV.



George R. Thoma is Chief of the Communications Engineering Branch of the Lister Hill National Center for Biomedical Communications, a research and development division of the U.S. National Library of Medicine at the National Institutes of Health. In this capacity, he directs intramural R&D in mission-critical projects such as the automated extraction of bibliographic data from medical articles to populate the MEDLINE database, a system to aid family reunification in mass disaster events, automated screening software for TB and malaria, and other projects relying on document image analysis and understanding, biomedical image processing and machine learning. He earned a B.S. degree from Swarthmore College, and an M.S. degree and Ph.D. from the University of Pennsylvania, all in Electrical Engineering. Dr. Thoma is a Fellow of the SPIE, the International Society for Optical Engineering.