



Binary handwriting image enhancement by directional field-guided morphology

Marcin Adamski^{a,*}, Kacper Sarnacki^b, Khalid Saeed^a

^a Faculty of Computer Science, Białystok University of Technology, ul. Wiejska 45A, 15-351 Białystok, Poland

^b Faculty of Mathematics and Information Science, Warsaw University of Technology, ul. Koszykowa 75, 00-662 Warszawa, Poland

ARTICLE INFO

Article history:

Received 9 April 2020

Received in revised form 26 August 2020

Accepted 12 November 2020

Available online 2 December 2020

Keywords:

Handwriting

Image Filtering

Binarisation

Guided morphology

Directional field

ABSTRACT

This paper proposes a technique for processing handwriting images. The algorithm used in this study is **an improvement to the binarisation process. The enhancement focuses on correcting damaged lines that usually arise during the binarisation process, particularly, spurious holes, discontinuities, and eroded boundaries.** The presented method uses a **morphological dilation operation** in which a structural element is **locally adapted** using the information from a **directional field. The adaptation process involves a new criterion for selecting orientation and shape of a structural element that combines directional field, a coherence measure, and a circular histogram.** The field was computed using **gradient-based approach**, and a method based on a Hessian matrix. During experiments, our method was applied to the output of selected binarisation algorithms. The experiments were conducted on grayscale signature images (from the CEDAR database) and handwriting images (from the DIBCO database). The results of the algorithm were compared to the results of standard morphological operations (dilation, erosion, opening, and closing) and median filtering. The experiments show that the proposed method achieves significant accuracy improvement (8%–12% for Acc, 15%–32% for Acc2 measures), reduces the number of unwanted artefacts, and produces images with less distortion compared to those from standard approaches.

© 2020 Elsevier Inc. All rights reserved.

1. Introduction

Studies on automatic processing of images of handwritten documents have a long history. Many complete systems have been proposed in scientific literature, and for commercial applications. For signatures, a special type of handwriting, the main task is verifying signature authenticity or identifying the signer [23]. The main process in handwriting recognition systems is usually composed of several stages, such as **preprocessing, feature extraction, and classification.** One of the most commonly used techniques at the preprocessing stage is **binarisation. It is responsible for distinguishing two classes of pixels: one representing the background, and the other the handwriting.** It is a crucial task for further stages. Preprocessing results affect the performance of the entire system. **Unfortunately, images obtained after the binarisation process usually contain errors, among which the most common are damaged lines or noise.** This can prevent the proper execution of subsequent operations, such as thinning or line tracking. Despite many years of research on binarisation algorithms, there is no perfect solution to this problem. The available methods struggle with different problems that become more significant,

* Corresponding author.

E-mail address: m.adamski@pb.edu.pl (M. Adamski).

especially in case of low-quality images [8,34]. For this reason, after binarisation, additional processing is often applied to remove unwanted artefacts. These operations include median filtering and morphological operations. The disadvantage of these operations is that, in addition to removing some of the artefacts, they create new errors (e.g., merging lines that are close to each other, filling loops created intentionally, or breaking the line's continuity). These errors are especially important for signatures.

Herein a new method of binarised image enhancement is presented. The method is a novel application of directional field (DF) guided morphology to filtering binarised images of handwriting documents. The process of structural elements (SE) selection is based on a new criterion utilising coherence measure and circular histogram and allows to handle multiple line directions in a single location. By selecting the shape and orientation of the element, it is possible to reduce the number of errors, such as incorrect joining or filling of loops. The proposed method is applied to the output of the binarisation process. That is, it postprocesses the output images after the binarisation stage. Moreover, it also uses input grayscale image to compute the directional field that guides the postprocessing.

The main objective of this study is to create a universal algorithm for improving the binarisation of handwritten documents and other types of images that contain objects with linear structures. The method should be applicable to any type of binarisation algorithm and allow to improve damaged lines whilst preserving separation of nearby structures and loops. The assessment should include both visual and objective evaluations based on accuracy measure and ground truth images.

2. Related works

Various filtering techniques can be found in works focused on typed and handwritten text binarisation [31]. For example, most of the methods submitted to the Competition on Document Image Binarisation (DIBCO) include filtering. These methods apply selective filtering before and after the binarisation step [34]. However, the algorithms are usually adapted to one particular binarisation, and the authors often do not provide any details on how to parameterise them.

In binarisation framework presented in [9] a combination of three techniques was used to postprocess binary image: anisotropic diffusion for speckle reduction, filling gaps using strokes extracted from feature points, and enhancing text regions by swell filter. The presented techniques were applied to the output of selected binarisation method – a modified version of Wellner's adaptive thresholding [9]. In [37] the morphological closing and opening operations were utilized to filter documents binarised by a method based on texture features extracted from the image filtered by the Gabor filter bank. The authors of [30] used connected components to detect and remove artefacts left after binarisation method based on image fusion and background subtraction.

A similar technique was also used in [19], where groups of 8-connected pixels with an area smaller than computed threshold were removed. A novel approach was introduced in [20] uses convolutional neural networks, which significantly improve the effects of Otsu binarisation. However, the algorithm also works for only one type of binarisation. Moreover, similar results can be achieved with a more universal algorithm, such as the one introduced here.

Whilst there are many studies on the various approaches to signature feature extraction and classification, there are not a lot of studies devoted to preprocessing of signature images [11,23]. To reduce artefacts in a binary image, most of the works employed a median filter [3,25], an average filter [22], or basic morphology operations: dilation, erosion, and their superposition, such as opening and closing [16,22]. Another type of preprocessing technique is based on removing “hairy” edges, which are a result of the binarisation procedure. In [15], such artefacts were eliminated by converting ink pixels on signature strokes to background pixels, if left and right or upper and lower pixels were ink pixels.

One might find similar approaches to image filtering [40], as proposed in this study, where spatially variant SEs were applied to grayscale and binary images. However, the authors of [40] used a different technique for DF estimation and selection of SEs. The DF was estimated using an average gradient followed by regularisation with vector flow. This method enabled the direction at every point of the picture to be calculated by a propagating gradient value from the edges, where it is well defined. The SE had a rectangular form, in which the ratio of the side lengths was selected based on the distance from the edge; with the element being linear when close to the edge and taking on a square form at a greater distance. However, this method is not well suited for images with handwritten script. Because of the narrow width of the text curve, gradient propagation from the edge is usually not necessary, and may pose problems when the image has low quality and noise. If the edge is not correctly determined, it may result in a wrong choice of SE and its orientation. Moreover, in [40], the authors did not include text images in their experiments, and the assessment was based only on visual inspection.

Extraction of line-like structures is also important in the field of line drawings vectorization from raster images. Whilst the aim of these methods is different, namely they are used to obtain succinct shape representation in terms of vectors and parametric curves such as Bezier's [13,14], they also face the problem of segmenting the drawing from the background. The segmentation is performed at the initial stage of processing and its quality is important for the correctness of subsequent vectorization. In [14] the segmentation is performed using fully convolutional network U-net that assigns each pixel to one of two classes. The network is trained using real and synthetic images. This is very similar technique to the binarisation methods which also use CNNs [20,42]. This approach may produce very good results when assessed on prepared test set with ground truth images. However, the drawback is the preparation of the training set that must be both large and representative with characteristics that will be also encountered during the network prediction phase. For example, the CNN that was trained using technical drawings may not give optimal results when applied to handwritten digits, which limits its usage.

Moreover, the loss function used in training such networks is computed on per-pixel basis, and hence, it may not aim to preserve line continuity. An example of such a problem is given in section 6.5. In [5] the initial background separation is the simple global thresholding and may introduce artifacts that contribute to suboptimal vectorization which must be manually corrected [5]. To extract the drawing line the authors of [13] use cross correlation with predefined templates followed by high pass filtering, median filtering and Otsu binarisation. As can be seen in the examples presented in [13], in the case of low quality pictures such approach may introduce breaks between connected segments.

3. Outline of the proposed method

As an input of the algorithm, we accept a grayscale image of handwriting. The first step is to execute the binarisation process. Using the proposed method, we do not assume any specific approach: Any available solution for this purpose can be chosen.

In parallel to the binarisation process, a directional field [39] (DF) is calculated based on a grayscale image. The method for computing a DF should enable us to determine the direction precisely and it should be resistant to noise existing in the input. In our work, we analysed two approaches for determining the DF.

The first one is the algorithm proposed in [21]. It is a gradient-based approach that was natively applied to fingerprint images. In addition, the reliability of each value from a DF is assessed by using the coherence measure given in [4].

The second method is a Hessian-based approach. It uses the idea introduced by Frangi [17] for filtering medical images with vessel-like structures. As it was introduced in [2], Frangi's measure is distinguished by good noise immunity.

The next step is the postprocessing of a binary image using morphological operations. Dilation with locally adapted SEs is applied. For each position in the binary image, the shape of the SE is selected based on the DF. In case of the gradient-based approach, the decision also takes into account the coherence measure. If the value of the coherence exceeds the selected threshold, a linear SE is used with an orientation based on the value of the DF. However, if the coherence measure is lower, it is assumed that the direction value is not reliable. In such case two approaches were tested: application of diamond-shaped SE, or further analysis based on circular histogram to detect multiple directions at a given location.

As the final step, additional filtering is performed using a 4-neighbourhood test. The process fills small artefacts that might be left after the morphological processing.

Fig. 1 illustrates the image processing flow in the proposed algorithm.

4. Description of techniques used

In this section we describe in detail the methods that were used in the proposed solution and in the experiments.

4.1. Binarisation

Binarisation is a process that converts the input image, usually in grayscale, to a binarised image. Its aim is segmentation of the object of interest from background. Studies show that many variants have been developed. A comprehensive review describing over 40 methods can be found in [38], and more recent advances are presented in [8,20,24,41,42]. One also may find methods for multi-level binarisation that partition image pixels into three or more categories [29].

The postprocessing technique proposed in this study has been tested with the following algorithms: Otsu [33], Sauvola [36], and Bradley [6]. The Otsu method [33] is one of the most popular methods of binarisation. The threshold value is based on optimisation of maximising intra-class variance between two classes of pixels that represent the foreground (object) and the background. In its basic form the Otsu method is global, a single threshold value is selected for the entire image. This may cause incorrect segmentation when image contains noise or when it is non-uniformly illuminated. In such cases one may apply Otsu with iterative partitioning [8], where the image is initially partitioned into sections with bimodal histogram. Sauvola [36] and Bradley [6] are local binarisation methods – a threshold value is computed separately for each pixel based on its neighbourhood. Such methods adapt to local image characteristics and are more resistant to non-uniform document illumination.

4.2. Morphology

Morphology operations can be used to filter artifacts or merge disconnected components [18]. In particular, the erosion operation may remove small structures or thin lines, and therefore, it may be used for filtering and removing objects that are small enough to be interpreted as noise. In contrast, dilation is useful for object enlargement. It can be used to remove small holes inside the object perimeter and reconnect structures that were unexpectedly separated during previous processing. The main drawback is that dilation enlarges the objects, while erosion shrinks them. Such behaviour is often unintended. Therefore, it is a common practice to combine them. Erosion followed by dilation is known as opening, while dilation followed by erosion is called closing. Those methods are expected to filter out noise, keeping the size of the objects unchanged.

Nevertheless, standard versions of opening and closing may cause splitting of lines that were previously merged or filling loops that are the important features of the object. This is often caused by the fact that the size and shape of the SE used by

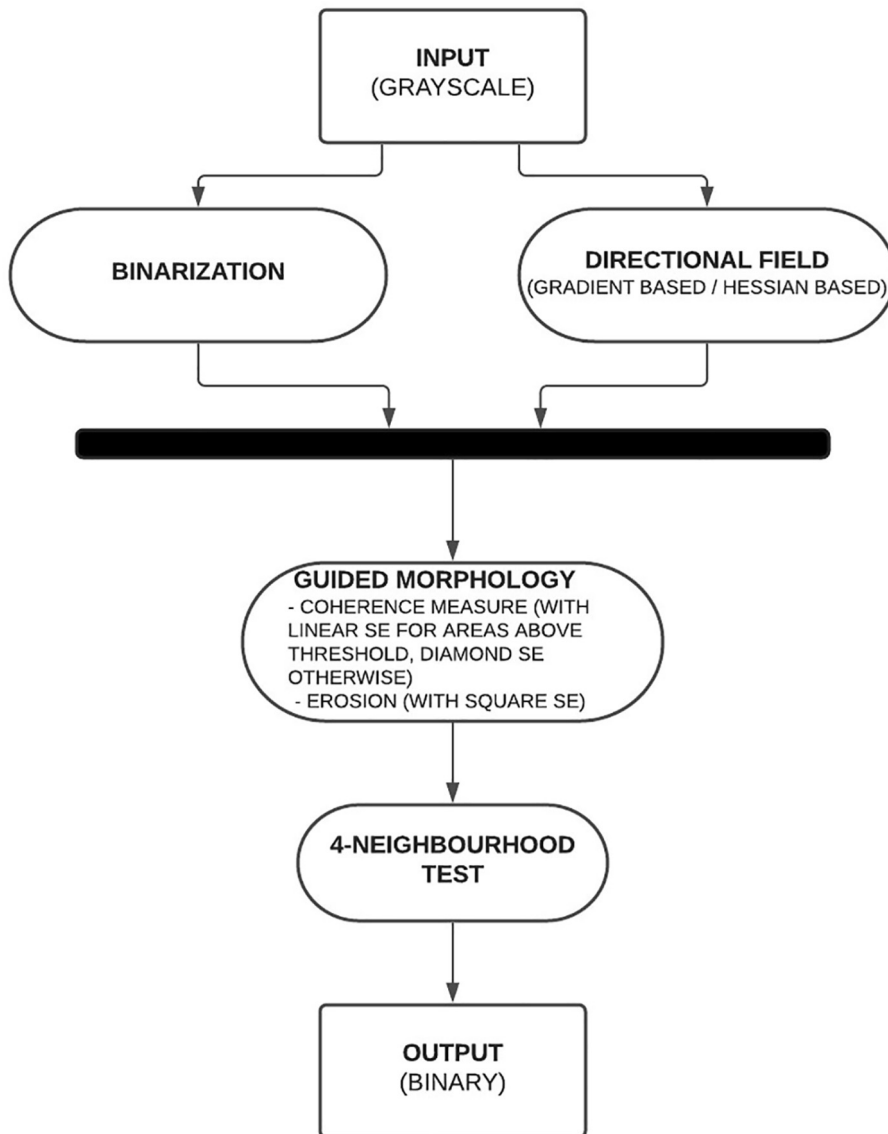


Fig. 1. Activity diagram of the proposed method.

morphological does not adapt to local characteristics of object in an image. To address this problem, in our work we propose a method for local selection of SE using information from directional field.

4.3. Directional field

The algorithm introduced uses a DF [39] to predict the orientation of the currently analysed line. Three approaches for solving the problem are presented below.

4.3.1. Gradient-based approach with a coherence measure

In this method, for each pixel of the input image (x,y) , we analyse its vicinity and we compute $\theta(x,y)$, the DF value of the point (information about the line orientation at this point). An exemplary image with a visualised DF is presented in Fig. 2(c). The image is divided into small segments, each of which contains a red line. The orientation of the line represents the dominant direction inside this segment. It can be observed that regions that are part of the handwriting line represent a direction that is consistent with the direction of the line in these regions. However, regions without any part of the line can be regarded as noise and hence ignored.

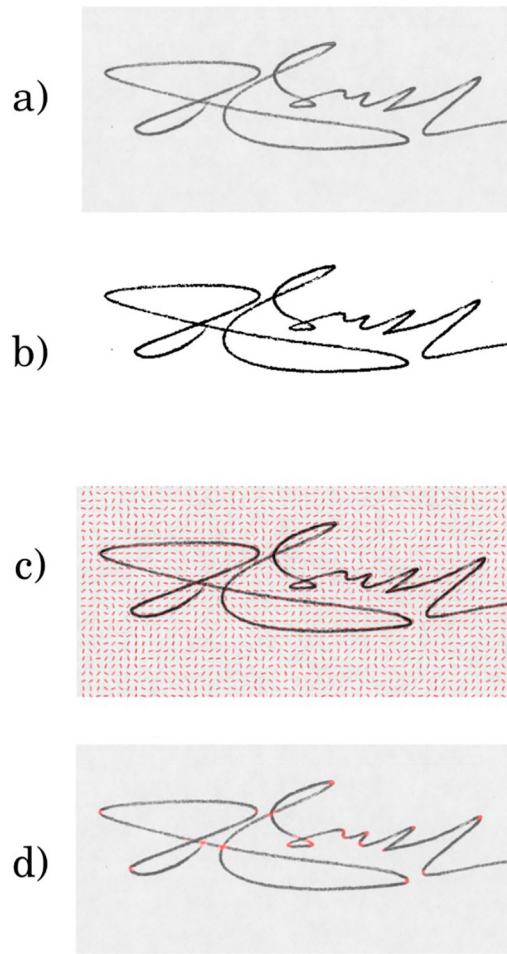


Fig. 2. Illustration of processing steps applied to a sample signature image from the CEDAR database. a) Input grayscale image. b) Binarised version of the image. c) Directional field computed for a grayscale signature image. Direction at a given point is shown as a short line segment. d) Illustration of the coherence measure. Areas with an intense red colour have the lowest coherence value.

There are multiple methods used for computing DFs. The most commonly used methods are based on **gradient operators** [28]. Having a grayscale image I , we compute a gradient along the x and y axes that will be defined as G_x and G_y , respectively:

$$\begin{bmatrix} G_x(x, y) \\ G_y(x, y) \end{bmatrix} = \begin{bmatrix} \frac{\partial I(x, y)}{\partial x} \\ \frac{\partial I(x, y)}{\partial y} \end{bmatrix} \quad (1)$$

There are various algorithms used to compute those gradients. One possible solution is to convolve the input image with a Prewitt or Sobel mask. We decided to use a filter based on a Gaussian filter along x and y directions that combines an initial smoothing with a gradient computation.

The gradients obtained for each pixel in the original image can be greatly affected by small changes in a local intensity caused by noise or other artefacts. To reduce rapid changes to the gradient resulting from factors such as noise, gradients can be averaged over a window with a specified size as given in [21].

When computing an average value, gradients in opposite directions should be added instead of being cancelled out, because they represent the same direction of a line. It is also beneficial to reduce the influence of gradients with low magnitude. These values usually represent regions with no evident edges.

One way to compute the average value is by treating the gradient vector as a complex number, and then computing the sum of the squared gradients G_W for each position (x, y) in the input image, with the window W centred at (x, y) :

$$G_W(x, y) = \sum_{(x', y') \in W} (G_x(x', y') + jG_y(x', y'))^2 \quad (2)$$

To smooth the change in a gradient direction, an additional low-pass filter can be applied. This step is achieved by normalising vectors G_W to unit length, and then by applying the Gaussian filter $K(x, y)$ as given in

$$G_S(x, y) = \frac{G_W(x, y)}{|G_W(x, y)|} \otimes K(x, y) \quad (3)$$

The final orientation of a line at a point (x, y) is obtained as an argument of G_S :

$$\theta(x, y) = \frac{1}{2} \arg(G_S(x, y)) + \frac{\pi}{2} \quad (4)$$

Because of the initial squaring, the angle must be divided by 2 and shifted by $\pi/2$ (the gradient being perpendicular to line boundary). The proposed method uses $\theta(x, y)$ to locally align the orientation of a linear SE.

In regions where there is no clearly edge defined, or the neighbourhood contains many edges in different directions (for example line intersection or dot-like shapes), the value from the DF is not correct. To rate the reliability of the DF, a ratio that measures the coherence of directions in the neighbourhood of a given position is defined as

$$r(x, y) = \frac{|\sum_{(x, y) \in W} G_S(x, y)|}{\sum_{(x, y) \in W} |G_S(x, y)|} \quad (5)$$

The value of this measure ranges from 0 to 1, where 1 denotes perfect coherence and 0 is a lack of coherence.

Fig. 2(d) illustrates this coefficient using the image of a sample signature. Areas with an intense red colour have the lowest coherence. As can be seen, these are areas with intersections or significant changes in the direction of the edges. In our algorithm, the value of $r(x, y)$ is compared with a heuristically selected threshold. If the value exceeds the threshold, a linear element is used at a position (x, y) during morphological postprocessing. The linear element is aligned according to $\theta(x, y)$. Otherwise, the direction given by $\theta(x, y)$ is considered to be unreliable and a diamond-shaped SE is used instead.

4.3.2. Gradient-based approach with a coherence measure and circular histogram

The main drawback of the previous method is that it does not adapt well to areas with intersections where there are two or more dominant directions. An example of such a situation is given in Fig. 3(a). For regions like those, a diamond-shaped SE is applied because of the low value of the coherence measure. Based on this observation, the following method is employed to identify all the directions.

The problem can be solved with the use of what is known as circular histogram. This is a common solution for problems based on the circular part of an image [32]. For this purpose, it is necessary to find the point where the lines are intersecting. The sliding window method is then used. For each fixed size window it is checked whether its central point is the point of

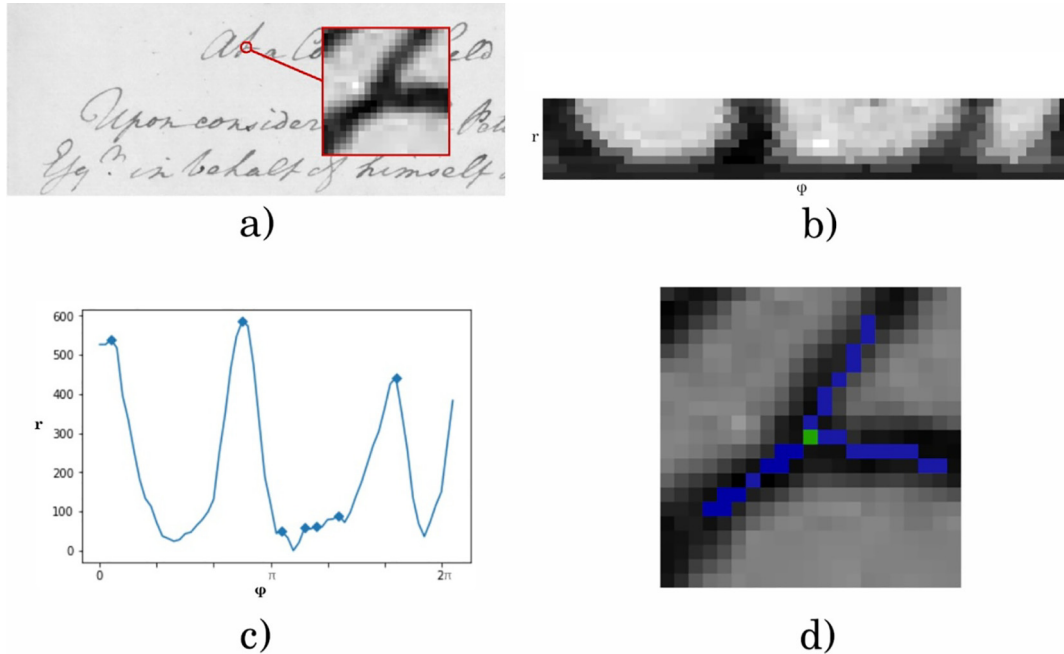


Fig. 3. Illustration of detecting multiple directions at a single location. a) Handwritten text image with an ambiguous area marked with a red square. b) The ambiguous area introduced in a) transformed from Cartesian into polar coordinates. c) Histogram of b) with all local maxima marked. d) Image of the ambiguous area with the centre point (green) and three branched lines detected by the algorithm (blue).

intersection. The intersection point is the point that connects at least three lines. How to identify a line from the centre point is shown in the next step.

The next step is to analyse each line extending from the centre point. This line can be represented by a tuple (φ, r) , where $\varphi \in [0, 2\pi]$ is the angle from the horizontal line extending from the centre and r is the length of the radius. Detailed values of the parameters that were used for the experiments are introduced in section 6.1.

This operation is computed by converting the area from polar into **Cartesian coordinates**. New coordinates (x_n, y_n) will be created with the use of

$$\begin{cases} x_n = r \cos \varphi + x_c \\ y_n = r \sin \varphi + y_c \end{cases} \quad (6)$$

where (x_c, y_c) is the centre of the currently analysed area, represented in Cartesian coordinates.

The results of such a transformation is shown in Fig. 3(b).

The last step is to construct the histogram function $h(\phi)$. For each angle ϕ and for specified r , the function computes the sum of the intensities of all pixels that construct a line starting at the centre point C of the area and having a length equal to r . It is assumed that the colour of the object has a higher intensity compared to the colour of the background. The exemplary results of the circular histogram are shown in Fig. 3(c). As can be observed, the function $h(\phi)$ contains local maxima. Three of the highest values represent the angle of the most distinctive lines of the analysed area.

To find the local maxima, it is necessary to analyse each point of the histogram and its neighbourhood. The sufficient condition for a local maximum in a point x_0 is to prove that, for a parameter $\delta > 0$, the following conditions are met:

$$f(x) \text{ is ascending for } x \in (x_0 - \delta, x_0) \quad (7)$$

$$f(x) \text{ is descending for } x \in (x_0, x_0 + \delta) \quad (8)$$

To detect only the proper directions in the histogram $h(\phi)$, it is necessary to define a threshold T . We used the average value of all the values of the histogram. All the maxima above the threshold are considered valid. Fig. 3(c) shows three valid maxima. The rest should be rejected. Fig. 3(d) shows the lines that were found by the algorithm.

As Fig. 3(d) shows, the three outstanding extrema represent all the lines that make up the intersection. Based on that, one can construct three linear SEs that are closely aligned with strokes going out of the intersection point. The SEs that will be created for the size of 5x5 are shown in Fig. 4.

4.3.3. Hessian-based approach

The third presented solution uses the idea **proposed in Frangi's algorithm [17]** for 2D images. In this case, instead of calculating a **DF that uses first derivatives, we find the Hessian and we compute its eigenvalues λ_1 and λ_2** .

The next step is to check whether the examined place is a background or a part of the line. We check it according to the table proposed by Frangi (Table 1).

We use the following measure proposed by Frangi for black ridges:

$$v_F = \begin{cases} 0, & \text{if } \lambda_2 < 0 \\ \exp\left(\frac{-R_B^2}{2\beta^2}\right) \left(1 - \exp\left(\frac{-S^2}{2c^2}\right)\right), & \text{otherwise} \end{cases} \quad (9)$$

where $R_B = \frac{\lambda_1}{\lambda_2}$, λ_1 and λ_2 are eigenvalues, with $|\lambda_1| < |\lambda_2|$, $\beta = 0.5$, $S = \sqrt{\lambda_1^2 + \lambda_2^2}$, and c is half the value of the maximum Hessian norm.

Determining whether the currently analysed place is noise involves examining the normalised coefficient v_F . If the currently analysed place is > 0.25 , it is assumed to be part of the line. The direction of the line is expected to be the same as an eigenvector of the smaller eigenvalue. The algorithm proceeds as follows:

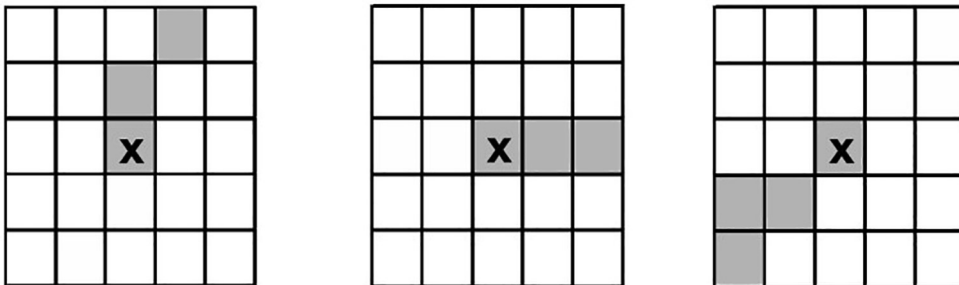


Fig. 4. Linear SEs adapted to the specific line intersection demonstrated in Fig. 3(d).

Table 1

Possible patterns depending on the value of the eigenvalues λ_k (where H = high, L = low, N = noisy (usually minor), and + or – indicates the sign of the eigenvalue). The eigenvalues are ordered as $|\lambda_1| \leq |\lambda_2|$.

2D		Orientation
λ_1	λ_2	
N	N	noisy, no preferred direction
L	H–	tubular structure (bright)
L	H+	tubular structure (dark)
H–	H–	blob-like structure (bright)
H+	H+	blob-like structure (dark)

1. Find the smaller eigenvalue.
2. Find an eigenvector that corresponds to the eigenvalue.
3. Determine the vertical or horizontal SE, depending on the direction of the eigenvector (analogously to the gradient-based approach introduced here).

4.4. Four-neighbourhood connectivity test

This algorithm is applied to remove some minor artefacts. Most of them are the size of one pixel. Therefore, they can be eliminated by applying the following connectivity test: The pixel is set to black when all pixels in its four-neighbourhood are also black.

4.5. Upsampling

To increase the number of possible directions to analyse, it is reasonable to try to upsample the image.

At the beginning of the algorithm, the original grayscale image is upsampled by 200% with the use of a cubic interpolation [26]. After this process, the height and width of the original image increases by a factor of 2. Analogously, the size of a SE increases by a factor of 2. Depending on a size of the SE, multiple lines can be constructed. For instance, for the size of 3×3 , there are exactly four possibilities: horizontal, vertical, and two diagonals. For the larger size of the SE, it is possible to create more lines. As previously described in our algorithm, this SE is used to represent a line more precisely. The choice of those lines depends on the angle of direction of a handwriting line in the currently analysed place.

The last step of the algorithm is downscaling, that is, returning the image to its original size. For every 4 pixels from the image, 1 pixel from the original image is assigned. The first step is to divide an upsampled image into 2×2 squares S . Every new pixel p_n is computed as a minimum from all the values of pixels of the corresponding square S_p :

$$\forall p_n p_n = \min(\forall_{p \in S_p} p) \quad (10)$$

5. Description of verification techniques used

5.1. Accuracy (Acc)

The accuracy of every image I_{test} was compared to the ideal (ground truth) image I_{ideal} according to the following formula:

$$S(I_{test}, I_{ideal}) = \frac{\sum_{(x,y) \in I_{ideal}} |sgn(I_{ideal}(x,y) - I_{test}(x,y))|}{\sum_{(x,y) \in I_{ideal}} 1} \cdot 100\% \quad (11)$$

The presented measure can be viewed as a classification of each pixel into one of two classes: object or background. Such approach will be useful when using statistical tools.

5.2. Improved accuracy (Acc2)

The testing method based on validating every pixel of the image has some disadvantages. Although it precisely compares similarity of the image processed by the algorithm to the ground truth image, background pixels and pixels of the text affect the measure to the same extent. Therefore, differences between tested algorithms may be small. We may be more interested in a measure that focuses on the pixels that make up the text and the area around the text.

To identify these pixels, binarisation with an average threshold was used, followed by dilation. Black pixels form a mask that contains the part of the image with key pixels. The results can be found in Fig. 5.



Fig. 5. Area included in improved accuracy measure: a) input image, b) the same image with area selected by accuracy measure marked in red colour.

For the exemplary image introduced in Fig. 5, the number of analysed pixels was reduced from 595,188 to 305,428 pixels, so it was reduced almost by a half.

The method gives results that are more sensitive to inaccuracies caused by line discontinuity or inappropriate thickness. Another benefit of the presented testing method is the fact that it includes also hard to process non-text elements. The most common elements are stains, folds of the paper or shadows.

5.3. Friedman test

The Friedman test is a statistical test that allows to verify whether the differences in accuracies of individual algorithms are statistically significant [10].

The idea of the algorithm is to rank k algorithms, each of which contains N tests. The evaluation consists in rating each algorithm: for the best performing algorithm the rank 1, the second best rank 2 etc. The final rank of every j -th algorithm is defined as follows:

$$R_j = \frac{1}{N} \sum_i r_i^j \quad (12)$$

Finally, the Friedman statistics is distributed according to the formula (13):

$$\chi_F^2 = \frac{12n}{k(k+1)} \left[\sum_j R_j^2 - \frac{k(k+1)^2}{4} \right] \quad (13)$$

The calculated value should be compared to the critical value obtained from χ^2 table. Critical value used for the computations is 5%. The number of degrees of freedom depends on the number of the analysed algorithms and equals $k - 1$.

If the calculated value is smaller compared to the critical value, null-hypothesis states that all the algorithms are equivalent and cannot be rejected.

5.4. Nemenyi test

Nemenyi test is a post-hoc test that can be proceeded if the null-hypothesis was rejected. The method allows to indicate which pairs of algorithms are equivalent [10].

The algorithms are compared pairwise. They are significantly different if the difference between their ranks is greater or equal to the critical difference (CD) value:

$$CD = q_\alpha \sqrt{\frac{k(k+1)}{6N}} \quad (14)$$

where $\alpha = 0.05$ and is based on the studentized range statistic divided by $\sqrt{2}$.

6. Tests and results

The tests were conducted on three databases: CEDAR [7], DIBCO (2009–2016) [12] and PBOK [1]. CEDAR contains images of handwritten signatures. This database was used to perform the initial evaluation of the results of the algorithm. The strengths and weaknesses of the algorithm could be determined. The evaluation enabled us to visualise how the algorithm handles broken lines and noise.

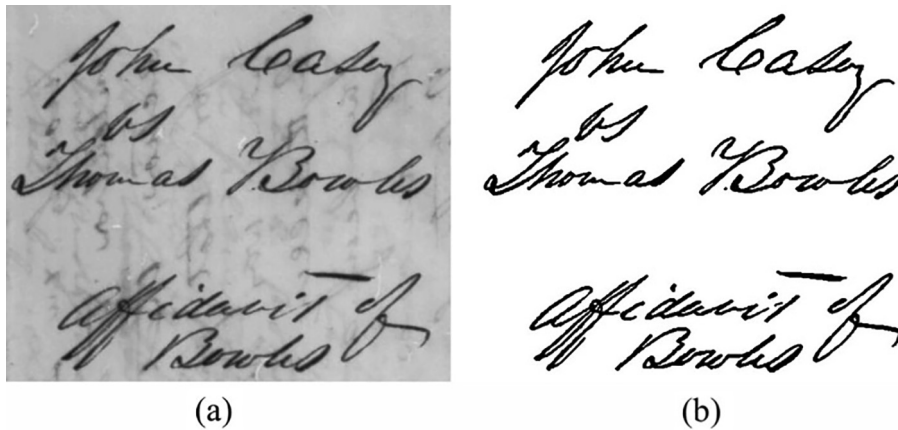


Fig. 6. Example images from the DIBCO database: unprocessed (a) and its ground truth binary version (b).

Forty-eight images were chosen from DIBCO [12], with seven images randomly chosen from each year (excluding 2015, for which there were no new images). The images contain ground truth versions, making it possible to define an accuracy measure. A sample pair of images before preprocessing and after ideal segmentation is shown in Fig. 6.

PBOK database contains various non-Latin handwriting images. One hundred images with Persian text were chosen from the database. The images, similarly to the DIBCO database, contain ground truth versions and will be used to evaluate accuracy of the algorithm.

6.1. Parameters adjustment

The presented algorithm is universal enough to be applied to various problems, in which thin objects identification is required. Depending on the problem, different parameters should be chosen. The following instructions should be used for selecting parameters for a handwriting text problem. For the other problems the described instructions may not give the best results.

The size of the structural elements depends on the size of the letter curvature, so it depends on the height of the letters. The best results are obtained with a mask of side 0.25 of the character height. In order to correctly fill the discontinuities of the line, it is important that the linear SE (causing dilation) is bigger than the square SE (causing erosion). The latter SE should be twice smaller.

For all of the tests, the parameters of the proposed method were the same. They have been selected based on a series of experiments and taking into account the above instructions:

- a 7×7 mask for computing the DF,
- a 15×15 mask for averaging gradients,
- a threshold value of 0.7 for a coherence coefficient,
- a 5×5 mask for a linear SE,
- a 5×5 mask for a diamond SE,
- a 3×3 mask for a square SE used during erosion,
- a radius of 10 for converting from polar to Cartesian coordinates during computation of the circular histogram, and
- a step ϕ of 0.1 during computation of the circular histogram.

6.2. Experiments I (CEDAR)

The following tests show the results of our algorithm with a gradient-based approach with a coherence measure used for computing the DF. It has been compared to the following standard postprocessing approaches:

- dilation (diamond),
- dilation (square),
- morphological close (diamond),
- morphological open (diamond), and
- median filtering.






















	Without post-processing	Morphology dilate(diamond)	Morphology dilate(square)	Morphology close(diamond)	Morphology open(diamond)	Median	Authors' method
Otsu							
Sauvola							
TBradley							

Fig. 7. Comparison of different postprocessing techniques applied to a threshold image from the CEDAR database. The rows represent different binarisation techniques. Their unprocessed outcome is given in the first column. Subsequent columns show the results of selected postprocessing methods; results from our approach are in the last column.

In all these methods, a 3×3 mask was used. Three different binarisation algorithms were used for comparison: the Otsu method, the Sauvola method and the Bradley method.

Our algorithm with the use of every binarisation method was compared with all the combinations of binarisation and postprocessing techniques. A visualisation of the results for an exemplary image is shown in Fig. 7.

As can be seen in Fig. 7, the main problem in this signature image is that the outputs of the binarisation algorithms contain artefacts along a signature line that need to be repaired. The aim of the correct postprocessing method is to improve a signature line without creating new errors. Morphological processing or median filtering enables removal of some of the artefacts. Nevertheless, new errors can also arise.

Fig. 8 presents magnified images of a line and loop sections that are largely affected by such defects. As illustrated in Fig. 8(a), the morphology open operation and median filtering may severely break line continuity if the binarised image contains too many defects. Line continuity is better preserved with filtering using dilate and close morphology operations. However, they can damage loops between parallel lines that are close to each other. This effect can be observed in Fig. 8(b). By using the algorithm proposed in this study, we were able to eliminate both of the problems, as can be seen in Fig. 8(a) and (b).

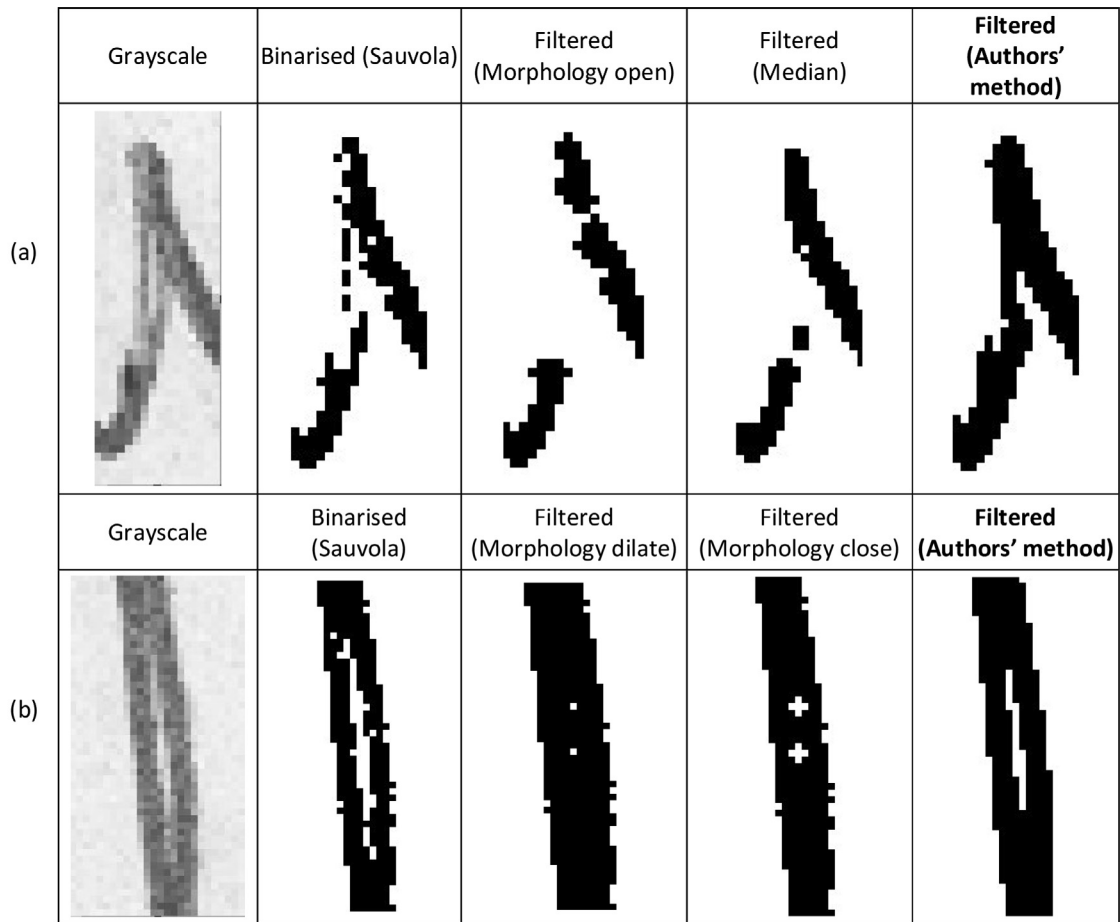


Fig. 8. Binarisation defects and results of filtering illustrated on magnified signature fragments: a) with cursive stroke, b) with loop element. The grayscale image is in the first column; its binarised version is given in the second column. Subsequent columns show results of selected filtering techniques.

Nevertheless, there are still cases when line damage cannot be removed. This may occur when the size of the damage exceeds the size of the SE. In the postprocessed image, damage such as large holes or a line discontinuity may appear. There are also cases in which a structure, such as a very small loop, may lead to a low coherence value of a DF in this region. As a result, it is incorrectly filled.

In the next part of the experiments, our research focused on further improvement of our method. A different way of computing the DF was examined and an application of upsampling was undertaken. For objective evaluation, ground truth images and an accuracy measure were used.

6.3. Experiments II (DIBCO)

The DIBCO and PBOK databases contain handwritten text. Because in this case we have access to ground truth images, it is possible to verify the obtained results with the expected values. Therefore, the accuracy measure introduced in sections 5.1 and 5.2 can be used.

The first test compares our gradient-based algorithm with a coherence measure (evaluated visually in the previous experiment) with various standard postprocessing methods. For each of these methods (dilation, opening, closing, and median filtering), a 5×5 SE was used.

Our algorithm uses dilation with SEs of different shapes. However, using dilation only thickens the objects, hence the use of closing is necessary. For this reason, erosion was used as the last step. A fixed square 3×3 pixel mask was applied.

The algorithms have been tested with different types of binarisation. The final results are listed in Table 2. The first row presents results of different types of binarisation for the same image. The other rows present different types of postprocessing, including our algorithm.

Table 2

Accuracy (Acc) and improved accuracy (Acc2) measure for selected binarisation and post-processing methods, compared with our algorithm.

		Otsu		Sauvola		Bradley	
		Acc [%]	Acc2 [%]	Acc [%]	Acc2 [%]	Acc [%]	Acc2 [%]
1	Without postprocessing	81.90	66.69	82.36	68.67	81.59	68.50
2	Morphology dilate (diamond)	83.94	73.47	83.35	68.74	83.45	69.01
3	Morphology dilate (square)	84.48	72.72	83.51	71.95	83.64	72.75
4	Morphology close (diamond)	83.57	71.59	85.31	74.62	85.72	77.08
5	Morphology open (diamond)	84.13	72.19	84.37	72.08	85.21	72.37
6	Median	84.76	74.21	84.68	72.58	85.56	75.28
7	Authors' method (gradient-based algorithm, coherence measure)	89.69	84.47	91.16	87.37	91.33	85.12
8	Authors' method (gradient-based algorithm, coherence measure, four-neighbourhood test)	90.06	85.20	91.19	86.12	92.53	87.65
9	Authors' method (gradient-based algorithm, coherence measure, circular histogram, four-neighbourhood test)	90.34	82.22	91.58	83.85	93.77	89.77
10	Authors' method (Hessian-based algorithm, four-neighbourhood test)	90.08	85.16	91.32	83.19	92.57	85.59
11	Authors' method (gradient-based algorithm, coherence measure, circular histogram, four-neighbourhood test, upscaling)	91.80	85.52	92.96	90.5	94.02	91.36

6.4. Statistical analysis of the results

Verification whether the differences of the accuracies of the algorithms are statistically significant will be carried out with the use of Friedman test. There are $k = 11$ groups (algorithms introduced in Table 2) and there were conducted $n = 148$ tests (number of images in the DIBCO and PBOC databases described in more detail in the introduction of section 6).

H_0 : Differences between accuracies of the algorithms 1–11 are not statistically significant.

H_A : Differences between accuracies of the algorithms 1–11 are statistically significant.

Average ranks of each algorithm are listed in Table 3.

The final value of Friedman statistics, according to the formula (13), equals:

$$\chi^2_F = 1315$$

For the used significance level and degrees of freedom ($d_f = 10$) critical value is 18,3070.

As $\chi^2_F > 18,3070$, null hypothesis is rejected and alternative hypothesis is supported.

The following test proved that differences between accuracies of the algorithms 1–11 are statistically significant. Next step is to find out if each pair of algorithms differ significantly. To this end, a post-hoc analysis with the use of Nemenyi test will be carried out.

Pairwise comparison between the algorithms is listed in Table 4. In line with the formula (14), the difference between algorithms is statistically significant if the value is greater than $CD = 1,24$.

The Nemenyi test proved that results of all versions of authors algorithm (7–11) are significant better than baseline approaches (1–6). The best performing version (11) turned out to have a significant advantage over all algorithms.

The experiments show that our method gives better results than all the other tested methods. It can also be observed that using only binarisation gives the worst results. Postprocessing algorithms significantly improve the results, what was proved with the use of the statistical tools.

Another test examined the effect of the four-neighbourhood test on the result of postprocessing in our algorithm. The accuracy of the algorithm with and without the connectivity test (seventh and eighth rows in Table 2) were calculated. The only aim of using the connectivity algorithm was to remove small artefacts (resulting from inaccurate binarisation). In fact, the connectivity algorithm slightly improves the results by 0.5%, what is not statistically significant improvement.

Next, the algorithm with a coherence measure and circular histogram was tested. As it adapts to the regions that are ambiguous for the previous gradient-based algorithm, it was expected to give better results. As illustrated in the ninth row of Table 2, the modification improves the accuracy of the gradient-based algorithm, as expected. For each type of binarisation, the algorithm allows for better imitation of the ground truth images.

The next tests concerned the Hessian-based version of our algorithm. The goal was to check whether the calculations on the second derivatives differ from the results of the algorithm that uses only the first derivatives (the gradient-based method). The average results of the tests of the algorithm based on the DIBCO database are listed in the tenth row of Table 2. The results show that this method gives results similar to those of the gradient-based approach. The biggest advantage that could be expected from the presented algorithm is its better resistance to noise than the gradient-based algorithm. Nonetheless, the presented tests show that it does not give a significant advantage compared to the gradient-based approach.

Table 3

Average rank of each algorithm 1–11 introduced in Table 2.

	1	2	3	4	5	6	7	8	9	10	11
R	10,68	9,29	9,18	6,63	7,84	7,39	3,76	3,67	3,00	3,33	1,24

Table 4

Pairwise comparison of statistical difference between accuracies of the algorithms 1–11 introduced in Table 2. Pairs with a statistically insignificant difference were marked gray.

	1	2	3	4	5	6	7	8	9	10	11
1	-	1,39	1,5	4,05	2,84	3,28	6,91	7,01	7,68	7,34	9,44
2	-	-	0,11	2,66	1,45	1,9	5,53	5,62	6,29	5,96	8,05
3	-	-	-	2,55	1,34	1,78	5,41	5,51	6,18	5,84	7,94
4	-	-	-	-	1,21	0,76	2,86	2,96	3,63	3,3	5,39
5	-	-	-	-	-	0,45	4,07	4,17	4,84	4,51	6,6
6	-	-	-	-	-	-	3,63	3,72	4,39	4,06	6,16
7	-	-	-	-	-	-	-	0,09	0,76	0,43	2,53
8	-	-	-	-	-	-	-	-	0,67	0,34	2,43
9	-	-	-	-	-	-	-	-	-	0,33	1,76
10	-	-	-	-	-	-	-	-	-	-	2,09
11	-	-	-	-	-	-	-	-	-	-	-

The last test was conducted with the use of upscaling. The algorithm was tested using various binarisation methods and a gradient-based approach. The average results of the accuracy measure are given in the last row of Table 2. The table shows that this version of our algorithm enables the best score to be achieved out of all the versions presented here. It can be observed that an increased number of analysed directions (through the use of more directional SEs) improves the image and helps to get an effect similar to that expected.

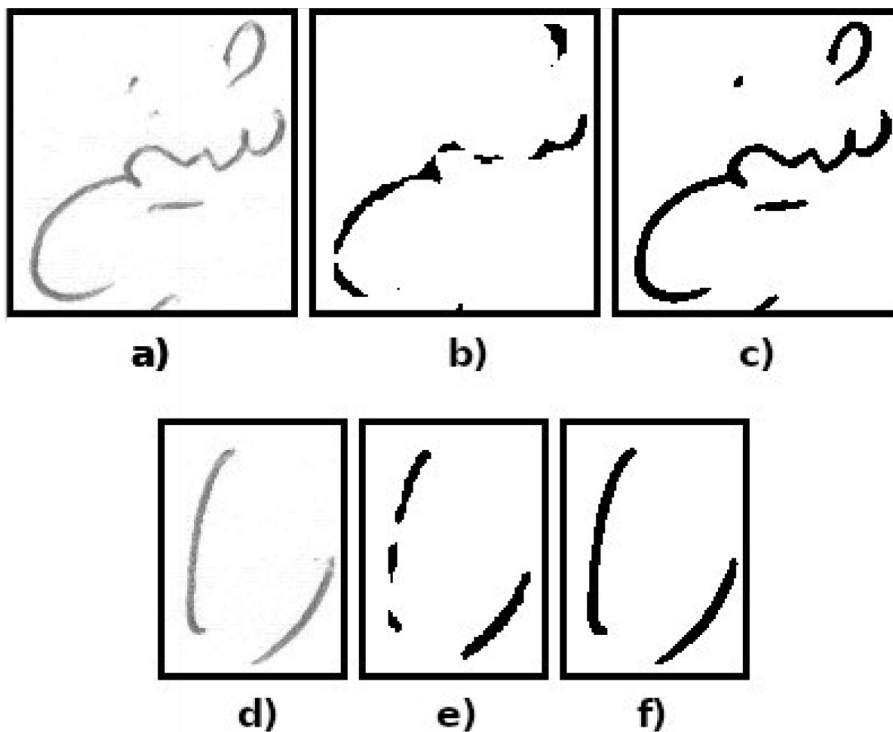


Fig. 9. Illustration of results of DSN binarisation and authors' method for specific glyphs of Arabic alphabet. a), d) grayscale glyphs. b), e) results DSN binarisation algorithm. c), f) results of authors' algorithm applied to the output of DSN binarisation.

6.5. Experiments with CNN-based binarisation method

DSN binarisation introduced by Quang Nhat et al. in [42] is an algorithm based on deep convolutional network. It was pre-trained with images from eight public datasets. The algorithm gives state-of-the-art results when assessed on images from DIBCO database. However, a closer examination shows that in certain cases this approach introduces line breaks even in relatively clean pictures. Fig. 9 shows exemplary images from an Arabic word as taken from PBOK dataset [1], where line continuity was not preserved. In this experiment we used the implementation of DSN binarisation provided by its authors in public repository [42].

As stated earlier, the authors' method proposed in this paper can be used with any binarisation algorithm. It can be seen in Fig. 9c and f, when applied to the result of state-of-the-art DSN binarisation method – its output is improved (Fig. 9b, Fig. 9c) by eliminating the smaller discontinuities. The experiments conducted on the whole dataset assessed quantitatively using accuracy measure did not show significant improvement. This was mostly due to the fact that the average result obtained from DSN on its own was already very high and our filtering method results in minor object inflation.

7. Conclusions

A new method for handwritten image enhancement has been presented. Various modifications of the algorithm were presented and tested on the basis of an established accuracy measure to the reference image.

The results of the tests indicate that the presented method positively affects the effects of binarisation, giving an image that is more similar to the expected, ideal result. In particular, the modification of the algorithm that doubles the image size turned out to give the best results. It may be surprising that the use of the Hessian, based on Frangi's method, does not significantly improve the results.

The algorithm enables a reduction in the number of unwanted artefacts created during the binarisation process, as well as the repair of destroyed lines, which is illustrated in Fig. 8(a). The method can be applied to the output of any binarisation technique provided the initial grayscale image is available to compute a DF. The results also demonstrate that the algorithm can be applied to different binarisation techniques to obtain better results with the same parameterisation.

The method used to determine the gradient map allows many directions to be specified at each point, making it is possible to process the intersections of lines correctly. Instead of choosing one dominant direction, or symmetrical SE (as diamond), every direction can be enhanced.

Applications of the presented authors' method are not restricted to handwriting. It may also be used in any task where image binarisation is applied to objects with linear structures, for ex: finger vein images analysed in biometrics [35] or printed circuit board images used in defects detection [27]. Author's approach can also be applied to some methodologies in their segmentation stage in drawing vectorization in order to reduce the number of interrupted segments in their final representation results.

CRedit authorship contribution statement

Marcin Adamski: Conceptualization, Methodology, Investigation, Software, Formal analysis, Validation, Writing - original draft, Writing - review & editing. **Kacper Sarnacki:** Investigation, Software, Data curation, Formal analysis, Validation, Visualization, Writing - review & editing. **Khalid Saeed:** Writing - review & editing, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by grants WZ/WI-IIT/3/2020, WZ/WI-IIT/4/2020 from Bialystok University of Technology and funded with resources for research by the Ministry of Science and Higher Education in Poland.

References

- [1] A. Alaei, U. Pal, P. Nagabhushan, Dataset and ground truth for handwritten text in four different scripts, *Int. J. Pattern Recognit Artif. Intell.* (2012).
- [2] M.A. Ansari, S. Zai, Y.S. Moon, Performance comparison of vesselness measures for segmentation of coronary arteries in 2D Angiograms, *Indian J. Sci. Technol.* 9(48) (2016).
- [3] R. Bajaj, S. Chaudhury, Signature verification using multiple neural classifiers, *Pattern Recogn.* 30 (1) (1997) 1–7.
- [4] A.M. Bazen, S.H. Gerez, Systematic methods for the computation of the directional fields and singular points of fingerprints, *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (7) (2002) 905–919.
- [5] M. Bessmeltsev, J. Solomon, Vectorization of line drawings via PolyVector FIELDS, *ACM Trans. Graphics* 38(1) (2018).
- [6] D. Bradley, G. Roth, Adaptive thresholding using the integral image, *J. Graphics Tools* 12 (2) (2007) 13–21.
- [7] CEDAR database (2007). <http://www.cedar.buffalo.edu/NIJ/data/signatures.rar> (accessed 15.03.2019)
- [8] N. Chaki, S.H. Shaikh, K. Saeed, *Exploring Image Binarization Techniques*, Springer Publishing Company, Incorporated, 2014.

- [9] Y. Chen, L. Wang, Broken and degraded document images binarization, *Neurocomputing* 237 (2017) 272–280.
- [10] J. Demsar, Statistical comparisons of classifiers over multiple data sets, *J. Mach. Learning Res.* 7 (2006) 1–30.
- [11] M. Diaz-Cabrera, A. Morales, M.A. Ferrer, Emerging issues for static handwritten signature biometrics, in: G. Pirlo, D. Impedovo, M. Fairhurst (Eds.), *Advances in Digital Handwritten Signature Processing*, World Scientific, 2014, pp. 109–120.
- [12] DIBCO database, 2017. <https://vc.ee.duth.gr/dibco2017> (accessed 15.03.2019)
- [13] L. Donati, S. Cesano, A. Prati, A complete hand-drawn sketch vectorization framework, *Multimedia Tools Appl.* 78 (14) (2019) 19083–19113.
- [14] V. Egiastian, O. Voynov, A. Artemov, D. Volkhonskiy, A. Safin, M. Taktasheva, D. Zorin, E. Burnaev, Deep vectorization of technical drawings, in: 16th European Conference on Computer Vision, 2020.
- [15] M.A. Ferrer, J.F. Vargas, A. Morales, A. Ordez, Robustness of offline signature verification based on gray level features, *IEEE Trans. Inf. Forensics Secur.* 7 (3) (2012) 966–977.
- [16] J. Fierrez-Aguilar, N. Alonso-Hermira, G. Moreno-Marquez, J. Ortega-Garcia, An off-line signature verification system based on fusion of local and global information, in: *Biometric Authentication*. vol. 3087/2004, ed: Springer Berlin/Heidelberg, 2004, pp. 295–306.
- [17] A. Frangi, W. Niessen, K. Vinc, M. Viergever, Multiscale vessel enhancement filtering, in: *Medical Image Computing and Computer Assisted Intervention MICCAI98 Lecture Notes in Computer Science* 1496/1998: 130, 1998.
- [18] R.C. Gonzalez, R.E. Woods, *Digital Image Processing*, 3rd ed., Prentice-Hall Inc, 2007.
- [19] M.R. Gupta, N.P. Jacobson, E.K. Garcia, OCR binarization and image pre-processing for searching historical documents, *Pattern Recogn.* 40 (2) (2007) 389–397.
- [20] S. He, L. Schomaker, DeepOtsu: Document enhancement and binarization using iterative deep learning, *Pattern Recogn.* (2019).
- [21] L. Hong, Y. Wan, A. Jain, Fingerprint image enhancement: algorithm and performance evaluation, *IEEE Trans. Pattern Anal. Mach. Intell.* 20 (8) (1998) 777–789.
- [22] K. Huang, H. Yan, Off-line signature verification based on geometric feature extraction and neural network classification, *Pattern Recogn.* 30 (1) (1997) 9–17.
- [23] D. Impedovo, G. Pirlo, Automatic signature verification: the state of the art, *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* 38 (5) (2008) 609–635.
- [24] F. Jia, C. Shi, K. He, C. Wang, B. Xiao, Degraded document image binarization using structural symmetry of strokes, *Pattern Recogn.* 74 (2018) 225–240.
- [25] D.J. Kennard, W.A. Barrett, T.W. Sederberg, Offline signature verification and forgery detection using a 2-D geometric warping approach, in: 21st International Conference on Pattern Recognition (ICPR), Tsukuba, Japan, 2012, pp. 3733–3736.
- [26] R. Keys, Cubic convolution interpolation for digital image processing, *IEEE Trans. Acoust. Speech Signal Process.* 26 (6) (1981) 1153–1160.
- [27] D. Li, S. Li, W. Yuan, Flexible printed circuit fracture detection based on hypothesis testing strategy, *IEEE Access* 8 (2020) 24457–24470.
- [28] D. Maltoni, D. Maio, A. Jain, S. Prabhakar, *Handbook of Fingerprint Recognition*, Springer-Verlag, London, 2009.
- [29] F.D. Martino, S. Sessa, PSO image thresholding on images compressed via fuzzy transforms, *Inf. Sci.* 506 (2020) 308–324.
- [30] N. Mitianoudis, N. Papamarkos, Multi-spectral document image binarization using image fusion and background subtraction techniques, in: *IEEE International Conference on Image Processing (ICIP)*, 2014, pp. 5172–5176.
- [31] H. Michalak, K. Okarma, Improvement of image binarization methods using image preprocessing with local entropy filtering for alphanumeric character recognition purposes, *Entropy* 21 (2019) 562.
- [32] A. Molder, O. Martens, T. Saar, Adaptively undersampled, circular histogram based image processing for rotation invariant coin detection, 2012.
- [33] N. Otsu, A threshold selection method from gray-level histograms, *IEEE Trans. Syst. Man Cybernet.* 9 (1) (1979) 62–66.
- [34] I. Pratikakis, K. Zagoris, G. Barlas, B. Gatos, ICDAR2017 Competition on Document Image Binarization (DIBCO 2017), in: in 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), 2017, pp. 1395–1403.
- [35] H. Qin, X. He, X. Yao, H. Li, Finger-vein verification based on the curvature in Radon space, *Expert Syst. Appl.* 82 (2017) 151–161.
- [36] J. Sauvola, M. Pietikinen, Adaptive document image binarization, *Pattern Recogn.* 33 (2) (2000) 225–236.
- [37] A. Sehad, Y. Chibani, R. Hedjam, M. Cheriet, Gabor filter-based texture for ancient degraded document image binarization, *Pattern Anal. Appl.* 22 (1) (2019) 1–22.
- [38] M. Sezgin, B. Sankur, Survey over image thresholding techniques and quantitative performance evaluation, *J. Electron. Imaging* 13 (1) (2004) 146–168.
- [39] A. Vaxman, M. Campen, O. Diamanti, D. Bommers, K. Hildebrandt, M. Ben-Chen, D. Panozzo, Directional field synthesis, design, and processing, *SIGGRAPH '17 Courses* (2017).
- [40] R. Verd-Monedero, J. Angulo, J. Serra, Anisotropic morphological filters with spatially-variant structuring elements based on image dependent gradient fields, *IEEE Trans. Image Process.* 20 (1) (2011) 200–212.
- [41] G.D. Vo, C. Park, Robust regression for image binarization under heavy noise and nonuniform background, *Pattern Recogn.* 81 (2018) 224–239.
- [42] Q.N. Vo, S.H. Kim, H.J. Yang, G. Lee, Binarization of degraded document images based on hierarchical deep supervised network, *Pattern Recogn.* 74 (2018) 568–586.