Accurate Detection of Dysmorphic Nuclei Using Dynamic Programming and Supervised Classification

Marlies Verschuuren[1], Jonas De Vylder[2], Hannes Catrysse[3], Joke Robijns[1], Wilfried Philips[2], Winnok H. De Vos[1,4]*

1 Department of Veterinary Sciences, University of Antwerp, Antwerp, Belgium, 2 Department of Telecommunication and Information Processing, IPI, iMinds, Ghent University, Ghent, Belgium, 3 Institute for Agricultural and Fisheries Research (ILVO), Melle, Belgium, 4 Department of Molecular Biotechnology, Ghent University, Ghent, Belgium

* Winnok.DeVos@UAntwerpen.be

Abstract

A vast array of pathologies is typified by the presence of nuclei with an abnormal morphology. Dysmorphic nuclear phenotypes feature dramatic size changes or foldings, but also entail much subtler deviations such as nuclear protrusions called blebs. Due to their unpredictable size, shape and intensity, dysmorphic nuclei are often not accurately detected in standard image analysis routines. To enable accurate detection of dysmorphic nuclei in confocal and widefield fluorescence microscopy images, we have developed an automated segmentation algorithm, called Blebbed Nuclei Detector (BleND), which relies on two-pass thresholding for initial nuclear contour detection, and an optimal path finding algorithm, based on dynamic programming, for refining these contours. Using a robust error metric, we show that our method matches manual segmentation in terms of precision and outperforms state-of-the-art nuclear segmentation methods. Its high performance allowed for building and integrating a robust classifier that recognizes dysmorphic nuclei with an accuracy above 95%. The combined segmentation-classification routine is bound to facilitate nucleus-based diagnostics and enable real-time recognition of dysmorphic nuclei in intelligent microscopy workflows.

Introduction

Nuclear shape changes are present in a broad range of pathologies. Depending on the origin and cell type, nuclei of cancer cells display strikingly different sizes and overt shape alterations such as grooves, folds or lobes, as compared to normal cells [1,2]. Numerous disorders also demonstrate subtler morphological aberrations such as invaginations or protrusions. These protrusions are often referred to as nuclear blebs and they are characteristic for diseases of the nuclear lamina, *i.e.*, laminopathies [3,4]. In various laminopathies, these blebs represent weak spots, which can sometimes rupture causing illegitimate exchange of nuclear and cytoplasmic proteins [5–8]. Bleb formation has also been observed in viral infections, where it is considered

to represent a correlate of nuclear entry and/or egress [9,10]. Despite a clear correlation with disease, not all nuclei in a cell culture display crevices or blebs, and since their formation is time-dependent, it is imperative that they can be automatically detected with high fidelity, preferably in a large number of cells.

In fluorescence microscopy, nuclei are usually labelled using a DNA binding fluorescent dye, which facilitates their segmentation. Many automated nuclear segmentation methods have been described that rely on such a counterstain, including intensity-based [11], active contour [12,13], graph cut [14,15], region growing/merging [16] and dynamic programming-based methods [17,18]. These algorithms often require prior knowledge on the location (dynamic programming), intensity (graph cut) or shape (region merging) of the objects in the image. Unfortunately, dysmorphic nuclei, and more specifically, nuclei with blebs, are typified by subtle shape alterations and lower intensities inside blebs, thereby presenting a difficulty to most existing nuclear segmentation algorithms. To resolve this, we have devised a segmentation method for the detection of dysmorphic nuclei, called BleND (Blebbed Nuclei Detector). It is based on a two-pass thresholding to identify the approximate contours of nuclei, and an optimal path finding algorithm to refine these contours. We have used the algorithm to segment nuclei from a variety of cell types, and we have validated it on a ground truth data set using an integrated error metric. Its high performance allowed for building a robust classifier that accurately discriminates dysmorphic from normal nuclei.

Methods

Image data sets

To optimize and benchmark the BleND algorithm, an image data set (widefield microscopy) from DAPI-counterstained human dermal fibroblasts from a compound progeroid syndrome patient (HDF-NCP) was used [19]. This dataset was chosen because it shows high variability in nuclear phenotypes, with both normal and blebbed nuclei being present in the same culture. Additionally, the algorithm was validated with images of other DAPI-counterstained cell types acquired with different imaging modalities: human dermal fibroblasts with a lethal laminopathy phenotype due to a nonsense Y259X homozygous null mutations in the *LMNA* gene (HDF-NULL) [20], which show extremely dysmorphic nuclei, often with an intensity gradient in the nuclear DAPI signal due to chromatin reorganisation (here referred to as chromatin ruffling [21]); human dermal fibroblasts from a Hutchinson—Gilford Progeria syndrome patient (HDF-HGPS, widefield microscopy) [19]; CRISPR/CAS9-genome edited *ZMPSTE24* knockout HeLa cells (HeLa-ZKO; point scanning confocal microscopy); genome-edited *LMNA* knockout human HT-1080 fibrosarcoma cells (HT-LKO, widefield microscopy) [8]; mouse primary hippocampal neurons (spinning disk confocal microscopy) [22].

Widefield images were acquired using a Nikon Ti fluorescence microscope equipped with an Andor DU-885 X-266 camera. Point scanning confocal images were acquired with a Nikon A1R system and spinning disk confocal images were acquired with a Perkin Elmer Ultraview system both mounted on a Nikon Ti microscope. Acquisitions were performed using either a 40x dry (NA = 1.0) objective, 40x oil (Plan Apo, NA = 1.30) objective, or 60x oil objective (Plan Apo VC, NA = 1.40).

Image processing

BleND was implemented as a Java plugin in the image processing software FIJI [23], a packaged version of ImageJ [24], and is freely available at https://github.com/VerschuurenM/BLEND. The general pipeline is depicted in Fig 1A. In brief, after pre-processing, an intensity-based segmentation (two-pass thresholding) of the pre-processed image allows identifying
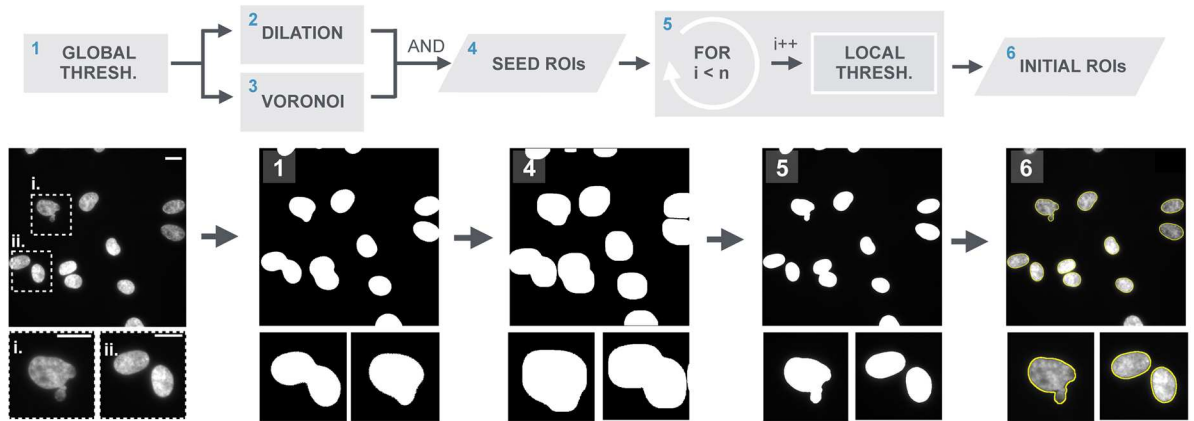
Fig 1. Overview of the BleND segmentation algorithm. (A) Workflow of the algorithm built on (B) intensity-based segmentation, (C) contour refinement, and (D) conditional watershed; (B) The segmentation process is implemented as a two-pass thresholding algorithm that generates "initial ROIs" of nuclei in the preprocessed image (i: dysmorphic nucleus, ii: two juxtaposed normal nuclei). A global thresholding is performed on the image, which creates a binary mask (1). The objects identified herein are dilated by 3 μm and combined (Boolean AND operation) with a Voronoi tessellation mask to ascertain that the dilated objects do not fuse. For each resulting "seed ROI" (4), a local threshold (5) is determined yielding an initial nuclear ROI (6) that is more accurate than the seed ROI (note the improved

segmentation for the dysmorphic and juxtaposed nuclei); (C) In the subsequent contour refinement procedure, the initial ROI is used (6) to straighten a 2μm wide region along the nuclear periphery (white dot indicates the point where the contour was opened and the white arrow indicates the direction of the straightening) (7). In this rectangular representation, the edge of the nucleus is enhanced by convolution with a vertical Sobel kernel (8). Then, an optimal path finding (OPF) algorithm determines the path with the highest path strength (9). The OPF algorithm effectively detects crevices surrounding nuclear blebs (red arrowhead). The contour of the nucleus is then reconstructed to generate a "refined ROI" and this process is repeated until the optimal path no longer changes (10); (D) To segment neighboring nuclei that could not be separated in the previous steps, a conditional watershed was implemented in which correct and incorrect splits were discriminated based on a size criterion and an intensity drop along the separation line (red arrowhead). This intensity drop is calculated as a median intensity profile perpendicular to the separation line (13). The user defines a threshold for the acquired intensity drop. In this example, the threshold is set at 0.75. If there is an intensity drop in the median profile of less than 25%—Min/Max intensity ratio above the 75% (dotted red) line (14)—the split is regarded as incorrect and the two parts of the nucleus are merged (15). If the drop is bigger, the split is regarded as being correct and it is retained to generate new nuclear ROIs.

doi:10.1371/journal.pone.0170688.g001

initial nuclear regions of interest (ROIs) in the image and generates contours that are refined using the contour refinement algorithm. Subsequently, adjacent nuclei are split using a conditional watershed algorithm. The contours (ROIs) that are newly generated in this process, will again be refined using the same contour refinement algorithm. The separate steps are described in more detail below.

Pre-processing. Background subtraction and multiple standard available linear and non-linear image filters (Gaussian, Median, Mean, Minimum, Maximum and Variance) are implemented in BleND; the scale of which can be defined by the user. This allows correcting for imperfect illumination, noise and intranuclear intensity variations (*e.g.*, chromocenters in mouse nuclei).

Segmentation. Since not all nuclei have the same average intensity, a global threshold can under- or overestimate their boundaries. In addition, blebs can have significantly lower intensities, causing them to become falsely assigned to the background (Fig 1B, inset i). To account for this problem, a two-pass thresholding was integrated that performs a rough global (image-based) thresholding, followed by a local (region-based) thresholding. Global thresholding serves to estimate the approximate location of all nuclei, whether they are clustered or not. The result is a set of "seed ROIs". These seed ROIs are conditionally dilated by maximally 3 μm, with their expansion being restricted by boundaries defined via Voronoi tessellation on the same seed ROIs. This prevents neighbouring regions from merging during the dilation process. Next, a local threshold is calculated within the conditionally dilated seed ROIs so as to obtain a better delineation of the actual contours (Fig 1B, inset i) and separation of neighbouring nuclei (Fig 1B, inset ii). The end result of two-pass thresholding is a set of "initial ROIs" for individual nuclei. All the automatic threshold algorithms that are implemented in FIJI were assessed for global and local thresholding (Huang [25], Intermodes [26], (IJ_)Isodata [27], Li [28], Maximum entropy [29], Mean [30], Minimum error [31], Minimum [26], Moment preserving [32], Otsu [33], Percentile [34], RenyiEntropy [29], Shanbhag [35], Triangle [36], and Yen [37]).

Contour refinement. After two-pass thresholding, crevices and invaginations surrounding blebs are not yet accurately delineated. In order to improve the initial ROIs, a contour refinement step was implemented (Fig 1C), which relies on contour straightening, a directional derivative and an optimal path finding algorithm. First, the boundary of the nucleus is straightened using an algorithm based on two-dimensional cubic splines [38], thereby generating a rectangular representation of a 2μm-wide region (1μm in both directions) surrounding the initial ROI (Fig 1C-7). Next, the edge of the nucleus is specifically enhanced by calculating the vertical derivative of the straightened image (Fig 1C-8). Finally, the exact contour is determined on the derivative image using an optimal path finding (OPF) algorithm (Fig 1C-9). Among all possible paths that can be drawn from left to right, the optimal path is found by maximizing the mean intensity of the path, defined as the ratio of the total intensity of the path

(gain) to the total path length (loss). We refer to this parameter as the "path strength". The underlying assumption is that the edge response (*i.e.*, the intensity of the derivative) will be the strongest at the true boundary of the nucleus. However, to prevent intranuclear intensity fluctuations or debris (also having a strong edge response) from skewing the boundary detection, a penalty is introduced for the total distance of the calculated path.
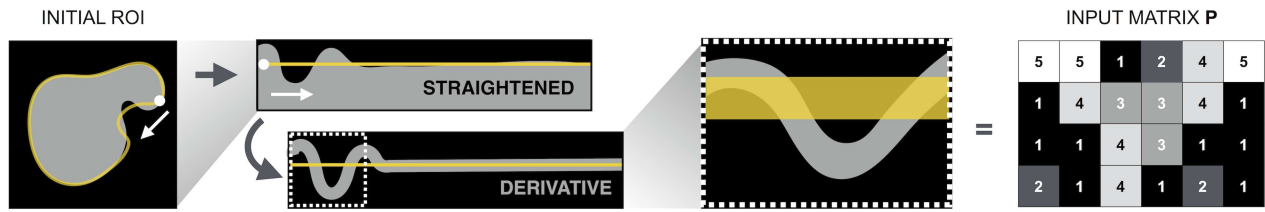
A numerical example of the OPF is represented in Fig 2. The derivative of the straightened image serves as input matrix **P** with dimensions (q, r), for the OPF algorithm (Fig 2A). The columns of the input matrix **P** are first divided by the column maxima (yielding normalized matrix **N**) to account for any declines in intensity that might occur in blebs or invaginations, so that they have an equal contribution to the average path strength. The optimal path is then calculated on the normalized matrix **N** using a dynamic programming approach. Starting from the left side of matrix **N**, the strength matrix **S**, gain matrix **G** and loss matrix **L** are simultaneously calculated (Fig 2B). Individual elements of each matrix (respectively $s_{i,j}$, $g_{i,j}$ and $l_{i,j}$) are recursively determined per column according to the strength function displayed in Eq 1. For all possible paths to element $n_{i,j}$ of **N**, a value for s is calculated, only to retain the path that provides the maximal path strength $s_{i,j}$ (Fig 2C). Since the elements of the preceding columns have already been determined, this procedure boils down to finding the optimal node (element with row index d) in the q rows of the former column (with index $j-1$). The corresponding gain ($g_{i,j}$) is determined by summing the value of this node ($g_{d,j-1}$) in matrix **G**, with the values of the elements of matrix **N** that lie in between $n_{d,j-1}$ and $n_{i,j}$. The corresponding loss ($l_{i,j}$) is determined by summing the value of this node ($l_{d,j-1}$) in matrix **L** with the number of matrix elements that lie in between $n_{d,j-1}$ and $n_{i,j}$ ($|n_{x,j-1}| + 1$ with x an element of [i,d[).

$$
\begin{cases}
s(N, G, L, i, j, k) = \dfrac{g_{k,j-1} + \sum_x n_{x,j-1} + n_{i,j}}{l_{k,j-1} + |n_{x,j-1}| + 1} \quad with\ x \in [\mathrm{i}, \mathrm{k}[\\[4mm]
d = \mathrm{argmax}_{k \in [1,q]}[s(N, G, L, i, j, k)] \\[4mm]
s_{i,j} = \dfrac{g_{i,j}}{l_{i,j}} = s(N, G, L, i, j, d) \\[4mm]
l_{i,j} = l_{d,j-1} + |n_{x,j-1}| + 1 \quad with\ x \in [\mathrm{i}, \mathrm{d}[\\[4mm]
g_{i,j} = s(N, G, L, i, j, d) \cdot l_{i,j} = g_{d,j-1} + \sum_x n_{x,j-1} + n_{i,j} \quad with\ x \in [\mathrm{i}, \mathrm{d}[
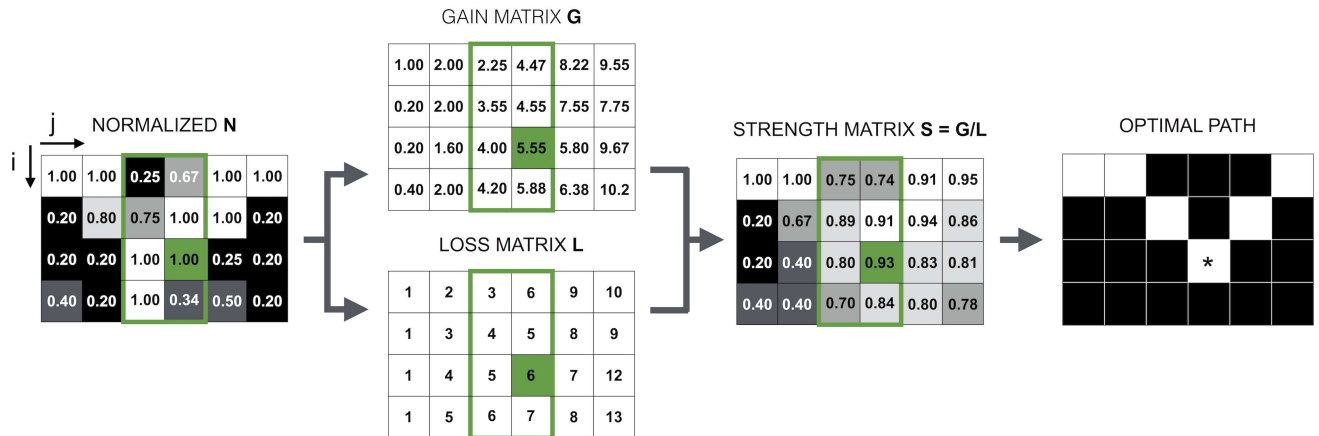\end{cases}
\tag{1}
$$

The optimal path is defined as the elements in each column of **S** with the highest value, allowing propagation angles up to 90° to accurately describe crevices surrounding blebs; this is in contrast with previously described methods where only angles < 45° were allowed [17,18] (Fig 1C-9). Once the optimal path is found on the derivative of the straightened image, it is converted to a closed contour, yielding the "refined ROI" for that specific nucleus (Fig 1C-10). The OPF is iteratively applied, each time using the newly created contour as substrate, for a defined number of cycles, to enable the detection of crevices that are bigger than 1μm, the half width of the rectangular region.

Conditional watershed. To separate clustered nuclei but prevent small structures like blebs from being disconnected, the watershed algorithm [39] was modified with two criteria for merging objects that were split incorrectly. The first criterion is based on object size: separate objects should not have an area below the minimal size assigned by the user. The second criterion is based on the presence of a sufficiently strong intensity decay (background signal) between adjacent nuclei, assuming that incorrectly split nuclei do not show this decay (Fig 1D) [8]. To robustly detect an intensity decay, a 3μm wide subregion is created around the separation line that arises from watershed segmentation (Fig 1D-13). For every pixel of the separation

A. INPUT GENERATION

INITIAL ROI

STRAIGHTENED

DERIVATIVE

INPUT MATRIX **P**

5	5	1	2	4	5
1	4	3	3	4	1
1	1	4	3	1	1
2	1	4	1	2	1

B. OPTIMAL PATH DERIVATION

GAIN MATRIX **G**

1.00	2.00	2.25	4.47	8.22	9.55
0.20	2.00	3.55	4.55	7.55	7.75
0.20	1.60	4.00	5.55	5.80	9.67
0.40	2.00	4.20	5.88	6.38	10.2

NORMALIZED **N**

j

i

1.00	1.00	0.25	0.67	1.00	1.00
0.20	0.80	0.75	1.00	1.00	0.20
0.20	0.20	1.00	1.00	0.25	0.20
0.40	0.20	1.00	0.34	0.50	0.20

LOSS MATRIX **L**

1	2	3	6	9	10
1	3	4	5	8	9
1	4	5	6	7	12
1	5	6	7	8	13

STRENGTH MATRIX **S = G/L**

1.00	1.00	0.75	0.74	0.91	0.95
0.20	0.67	0.89	0.91	0.94	0.86
0.20	0.40	0.80	0.93	0.83	0.81
0.40	0.40	0.70	0.84	0.80	0.78

OPTIMAL PATH

*

C. STRENGTH CALCULATION FOR ELEMENT i=3, j=4

k=1
$g_{3,4} = g_{1,3} + n_{2,3} + n_{3,3} + n_{3,4}$
$= 2.25 + 0.75 + 1 + 1 = 5.00$
$l_{3,4} = l_{1,3} + 3\ steps$
$= 3 + 3 = 6$
$s_{3,4} = g_{3,4} / l_{3,4}$
$= 5.00 / 6 = 0.83$

k=2
$g_{3,4} = g_{2,3} + n_{3,3} + n_{3,4}$
$= 3.55 + 1 + 1 = 5.55$
$l_{3,4} = l_{2,3} + 2\ steps$
$= 4 + 2 = 6$
$s_{3,4} = g_{3,4} / l_{3,4}$
$= 5.55 / 6 = $ **0.93**

OPTIMUM

k=3
$g_{3,4} = g_{3,3} + n_{3,4}$
$= 4 + 1 = 5.00$
$l_{3,4} = l_{3,3} + 1\ step$
$= 5 + 1 = 6$
$s_{3,4} = g_{3,4} / l_{3,4}$
$= 5.00 / 6 = 0.83$

k=4
$g_{3,4} = g_{4,3} + n_{3,4} + n_{3,3}$
$= 4.20 + 1 + 1 = 6.20$
$l_{3,4} = l_{4,3} + 2\ steps$
$= 6 + 2 = 8$
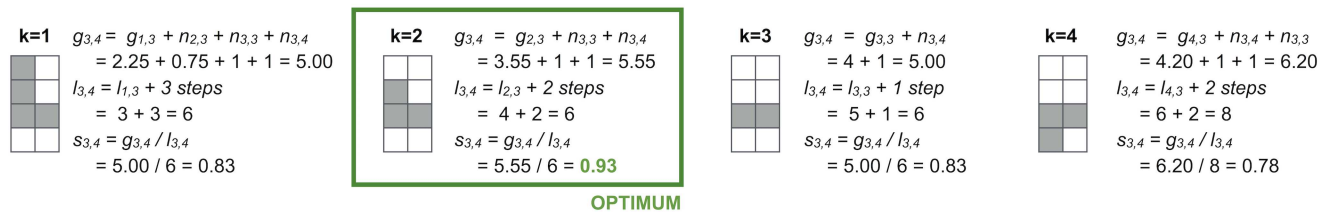$s_{3,4} = g_{3,4} / l_{3,4}$
$= 6.20 / 8 = 0.78$

Fig 2. Numerical example of the optimal path finding algorithm. (A) Schematic representation of the generation of input matrix **P**. The vertical derivative of the straightened representation of the nuclear periphery—defined as 2µm wide band surrounding the initial ROI—serves as input matrix **P** for the optimal path finding algorithm (OPF); (B) Columns of **P** are first normalized (divided by the resp. maximum) after which the optimal path is calculated on the normalized matrix **N** using a dynamic programming approach. Starting from the left side of the matrix, a strength function $s_{i,j}$ (strength matrix **S**) is calculated for every matrix element $n_{i,j}$ that takes the ratio of the sum of the intensity along the path ($g_{i,j}$; gain matrix **G**) and the total path length ($l_{i,j}$; loss matrix **L**). The optimal path in **S** is defined by the elements with the highest value per column; (C) Optimal path calculation for the element on row 3 and column 4 of matrix **N** ($n_{3,4}$, marked green in B). The algorithm calculates the gain ($g_{3,4}$), loss ($l_{3,4}$) and strength ($s_{3,4}$ = $g_{3,4}/l_{3,4}$) for all 4 possible paths starting in the previous column—in this case column 3—to element $n_{3,4}$. Since the algorithm is progressive, the values for gain, loss and strength have already been calculated for all elements in column 3. Thus, for element $n_{3,4}$, calculation of the gain comes down to summing the value of the first element k of the path in the gain matrix ($g_{k,3}$, the total gain up to that point) with the values of **N** along the rest of the path. Likewise, the loss for element $n_{3,4}$ is calculated by summing the value of the first element k of the path in the loss matrix ($l_{k,3}$, the total loss up to that point) with the total number of steps to $n_{3,4}$.

doi:10.1371/journal.pone.0170688.g002

line, a perpendicular intensity profile is measured along the width of the subregion, and the median of these individual intensity profiles is calculated ([Fig 1C-14](#)). If the min/max ratio of the median intensity profile is larger than a user-assigned cut-off (typically set at 75%), the split is regarded as incorrect and the two parts of the nucleus are merged. If the min/max ratio

is smaller, the split is regarded as being correct and two new nuclear ROIs are generated (Fig 1C-15).

Validation

Validation of the BleND segmentation algorithm was done by comparing the automatically detected contours (C) with manual delineations of 104 nuclei obtained from three independent observers (ground truth, GT_k, k = {1,2,3}). To quantify the segmentation performance, two error metrics were used: the average Hausdorff distance (AHD) and non-similarity index (NSI). The AHD is a proxy for the minimal distance between the automatically detected contours C and manually delineated contour GT (Eq 2) [40] and is calculated as follows: for all p points of contour C describing nucleus i, the minimum Euclidian distance (d) to contour GT_k is calculated. The average of these distances is the AHD for nuclei i with contour C as reference ($h(C_i, GT_{k,i})$). Since $h(C_i, GT_{k,i})$ is not equal to $h(GT_{k,i}, C_i)$, both are calculated and the maximum of these two values is retained as AHD between contour C_i and the kth ground truth for nucleus i. The NSI is calculated as the ratio of the non-overlapping area and the sum of the total area enclosed by both contours (C and GT, Eq 3) (Fig 3) [41]. For every nucleus i, both error metrics are scaled to a positive control (PC, Eq 4), which is defined as the average error of pairwise comparisons between the three independent GTs (Eq 5). The global error that was used to quantify the actual precision of the automated segmentation is the mean of both scaled parameters (Eq 6).

$$\begin{cases} h(C_i, GT_{k,i}) = \dfrac{1}{p} \sum_{a \in C_i} \min_{b \in GT_{k,i}} d(a, b) \\ \text{AHD}_i = \dfrac{1}{3} \sum_{k \in \{1,2,3\}} \max[h(C_i, GT_{k,i}), h(GT_{k,i}, C_i)] \end{cases} \quad (2)$$

$$\text{NSI}_i = \frac{1}{3} \sum_{k \in \{1,2,3\}} 1 - \frac{\text{Area}(C \wedge GT_{k,i})}{\text{Area}(C_i) + \text{Area}(GT_{k,i})} \quad (3)$$

$$\begin{cases} \text{AHD}_{i,scale} = \dfrac{\text{AHD}_i}{\text{AHD}_{i,PC}} \\ \text{NSI}_{i,scale} = \dfrac{\text{NSI}_i}{\text{NSI}_{i,PC}} \end{cases} \quad (4)$$

$$\begin{cases} \text{AHD}_{i,PC} = \dfrac{1}{6} \sum_{k \in \{1,2,3\}} \sum_{l \in \{1,2,3\}} \max[h(GT_{k,i}, GT_{l,i}), h(GT_{l,i}, GT_{k,i})] \\ \text{NSI}_{i,PC} = \dfrac{1}{6} \sum_{k \in \{1,2,3\}} \sum_{l \in \{1,2,3\}} 1 - \dfrac{\text{Area}(GT_{k,i} \wedge GT_{l,i})}{\text{Area}(GT_{k,i}) + \text{Area}(GT_{l,i})} \end{cases} \quad (5)$$

$$Error_i = \frac{1}{2}(\text{NSI}_{i,scale} + \text{AHD}_{i,scale}) \quad (6)$$

Data analysis

Classification of the nuclei, identified by the segmentation algorithm, was achieved using a morpho-textural feature set, including all standard ImageJ/FIJI shape attributes (area, perimeter, descriptors of the fitted ellipse, circularity, solidity) as well as curvature and texture
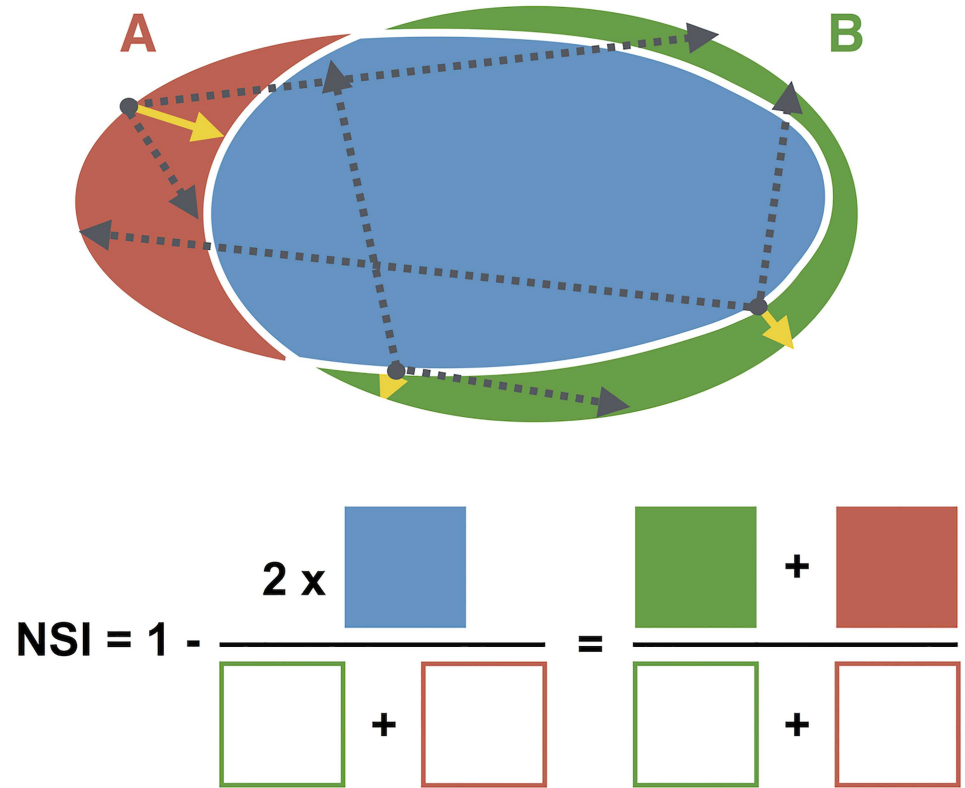
Fig 3. Schematic representation of the error metrics used for validation of the segmentation algorithm. Individual BleND segmentations (red line) were compared to the respective GTs (green line) using two error metrics: the average Hausdorff distance (AHD) and a non-similarity index (NSI). The AHD is calculated as the average of the minimal distances (yellow arrows)–selected among all possible distances (examples in dotted grey arrows)–between the pixels of both contours. The NSI is derived as the non-overlapping area (red and green area) divided by the sum of the total area described by these contours (red and green line).

descriptors. The curvature of the nuclear boundary (defined by the refined ROI) was depicted as the alteration of the orientation of subsequent edge segments. The total curvature was then calculated as the summation of the absolute values of the first derivatives of these segments. Other features describing the shape of the nuclei are the rotation-invariant elliptic Fourier descriptors (EFD) [42]. The gray-level co-occurrence matrix (GLCM) was calculated to extract features describing the texture of the cell nuclei [43]. The GLCM attributes obtained under different angles (0˚, 45˚, 90˚ and 135˚) were averaged to obtain rotation invariant parameters describing the texture.

The HDF-NCP data set consisted of 162 dysmorphic nuclei and 831 normal nuclei from which 162 were randomly sampled to obtain a dataset with equal fractions for the two classes.

Data analysis was done in R [44]. To explore the data and select the most informative features, principal component analysis (PCA) was performed. Selected features were used for training supervised classification schemes based on linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), regularized discriminant analysis (RDA), mixture discriminant analysis (MDA), naive Bayes (NB), flexible discriminant analysis (FDA), support vector machine (SVM), bagging (BAG), boosting (BOO) and random forest (RF). Different kernels were used in NB and SVM classifiers, whereas different regression methods were used to train FDA classification schemes. For classification, the dataset was split up in a test set (1/3) and training set (2/3) with equal class ratios. Using 10-fold cross-validation in the training set, the classification algorithms were trained and then used on the test set to determine misclassification rate (MCR) and false negative rate (FNR).

Results

Automatic segmentation matches manual delineation

To quantify the accuracy of the detection algorithm, the method was compared to three independent GTs of manually delineated HDF-NCP nuclei, using an integrated performance error, based on the average of AHD and NSI as described in the M&M section. The inter-individual variability of the GTs—calculated as the standard deviation of the error scores obtained after pairwise comparison of all GTs—was 7% for all nuclei and 9% for the dysmorphic nuclei only. Using the same error metric, a quantitative comparison was made of the segmentation algorithm using single (n = 16 threshold methods) or two-pass (n = 16 x 16 = 256 threshold combinations) thresholding, with and without contour refinement (Fig 4). On the complete dataset (Fig 4A), a single threshold could not attain the precision of the manual segmentation, since all of the obtained error scores were higher than those observed when comparing the GTs (grey dots). Two-pass thresholding improved the segmentation and resulted in 3 threshold combinations that lied within the GT error range (*i.e.*, the range of errors obtained by pairwise comparison of individual GT's, green-coded dots). Contour refinement boosted the performance and reduced the error scores significantly, resulting in 100 threshold combinations (39% of the 256 combinations) that resulted in values within this error range. The fact that multiple threshold combinations yielded errors within the GT error range indicates that the BleND algorithm attains the precision of manual delineation. For segmentation of dysmorphic nuclei, both 2-pass thresholding (P = 0.04756; Mann—Whitney U test, one-sided) and contour refinement (P = 0.004939; Mann—Whitney U test, one-sided) enhanced segmentation error scores in a statistically significant manner (Fig 4B). Some threshold methods served better as global method (e.g. Max Entropy) or as local method (e.g. Li), whereas others performed poorly throughout (e.g. Shanbhag).

BleND accurately delineates nuclei in a variety of cell types and data sets

To assess the generic value of the segmentation algorithm, BleND was also tested on images of a variety of cell types with aberrant nuclei such as HDF-HGPS, HDF-NULL, HeLa-ZKO, HT-LKO and mouse primary hippocampal neurons acquired with different imaging modalities at 40x or 60x magnification (Fig 5A). Lower magnifications were not considered since the main objective of BleND is to detect subtle deviations of nuclear shapes, which are not clearly visible at low resolution. All cells are characterised by dysmorphic nuclei, but they differ strongly in shape and texture: nuclei of HDF-NULL and HT-LKO cells are severely deformed and often show an intensity gradient due to chromatin ruffling, whereas HeLa-ZKO cells and
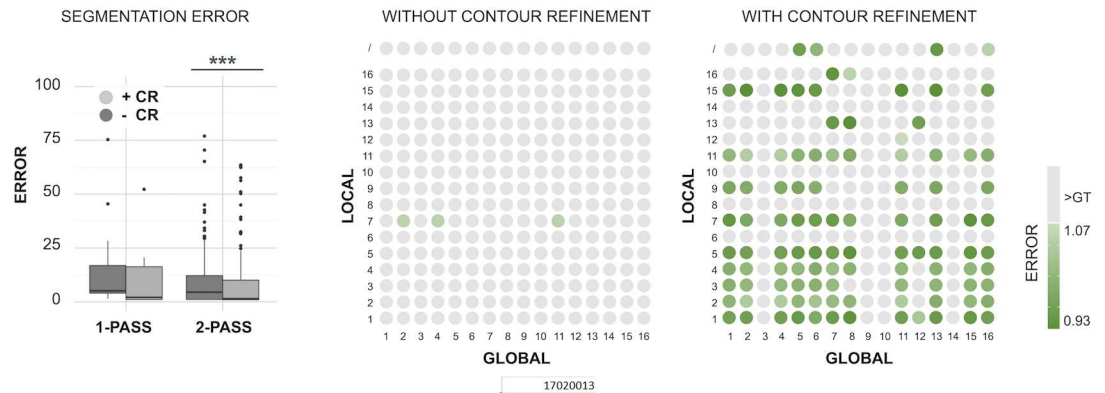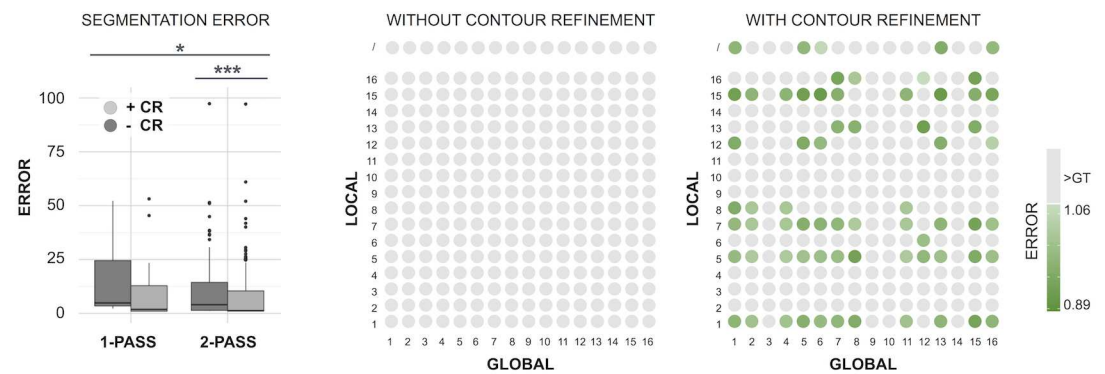
Fig 4. Comparison of performance errors for dysmorphic nuclei. Boxplots and dot plots of the performance errors of all threshold combinations using 1-pass (only global) or 2-pass thresholding (global and local), with (+CR) or without (-CR) contour refinement for (A) all nuclei, and (B) for dysmorphic nuclei only. Asterisks mark statistically significant differences according to the Wilcoxon rank-sum test (one sided) (* P < 0.05, *** P < 0.005). The outliers in the boxplots represent inadequate segmentations caused by an error-prone thresholding method. The color in the dot plots represents the error, with values falling within the error range of the ground truth (GT) comparisons displayed in green hues, and values exceeding this range in light grey. The numbers on the axes of the dot plots represent different threshold methods: 1: Huang, 2: Intermodes, 3: IJ_Isodata, 4: Isodata, 5: Li, 6: Maximum entropy, 7: Mean, 8: Minimum error, 9: Minimum, 10: Moment preserving, 11: Otsu, 12: Percentile, 13: RenyiEntropy, 14: Shanbhag, 15: Triangle, 16: Yen.

doi:10.1371/journal.pone.0170688.g004

primary hippocampal neurons have nuclei with small blebs and especially the latter show high intranuclear intensity spots (chromocenters). For all image types, suitable segmentation settings could be defined (Fig 5A).

Comparison with nuclear segmentation methods

Next, error scores of the BleND algorithm were compared with other state-of-the-art algorithms used for the analysis of nuclear morphology [12,45]. For this, a rough segmentation was performed using the 2-pass thresholding algorithm implemented in BleND, after which different refinement steps were compared: dynamic programming (BleND), level set (LS) active contour (Fiji) [23,45] and gradient vector flow (GVF) active contour (Matlab®) [12,13,46]. To allow accurate delineations of blebs and crevices, curvature penalty weights were decreased for both LS and GVF active contour. The results are represented in Table 1. GVF active contour and BleND showed a similar error and generated segmentations that attain the precision of the GT for normal nuclei. On the other hand, for dysmorphic nuclei,

A. COMPARISON OF CELL TYPES
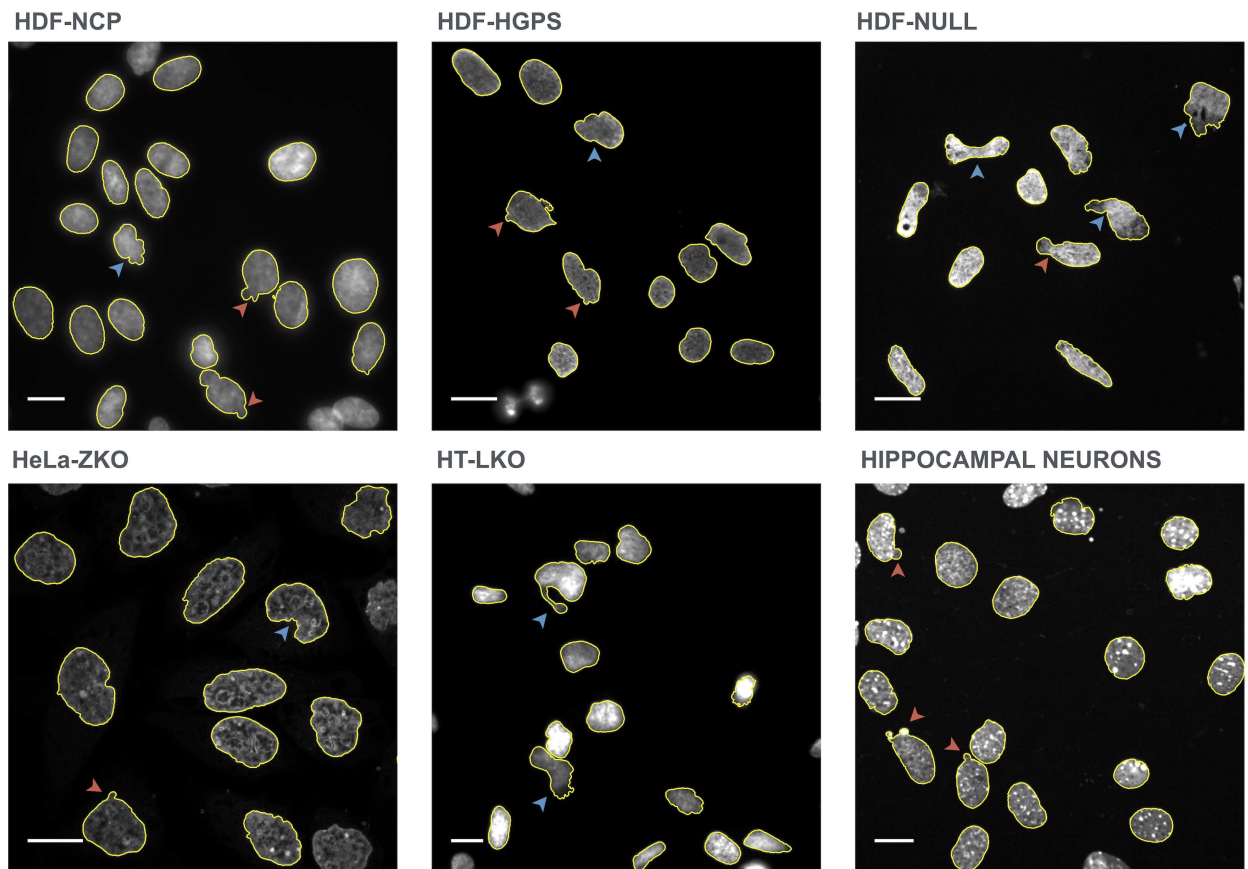
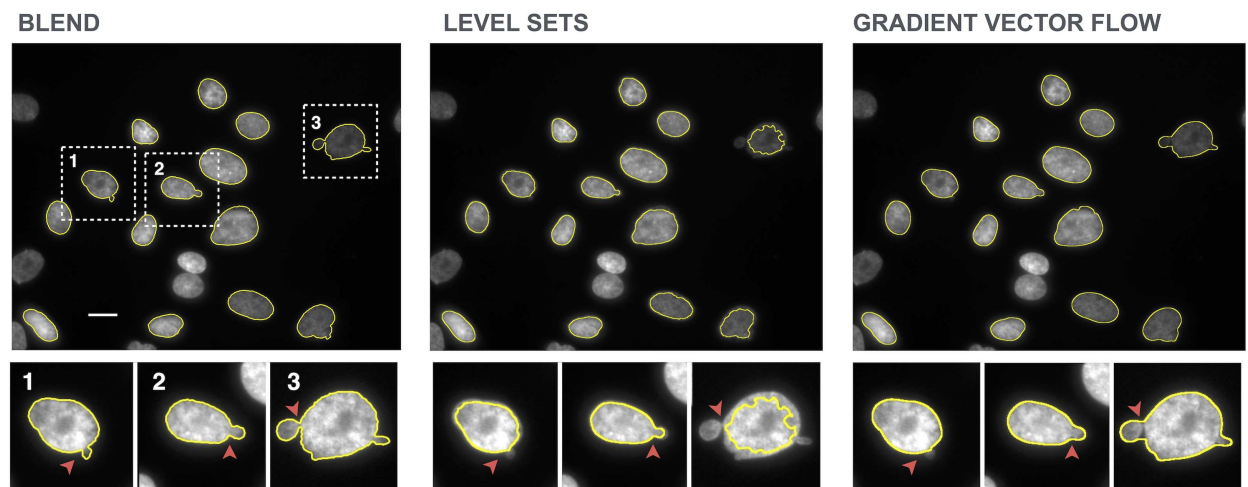## B. COMPARISON OF ALGORITHMS

Fig 5. Segmentation results for different cell types and contour refinement algorithms. (A) Segmentation results for DAPI counterstained nuclei of HDF-NCP (widefield microscopy), HDF-HGPS (widefield microscopy), HDF-NULL (widefield microscopy), HeLa-ZKO (point scanning confocal microscopy), HT-LKO (widefield microscopy) cells and mouse primary hippocampal neurons (spinning disk confocal microscopy). Blebbed (red arrowheads) and/or severely deformed nuclei (blue arrowheads) are accurately delineated; (B) Comparison of BleND with level set active contour and gradient vector flow active contour algorithms on an image of HDF-NCP cells. Insets show contrast-stretched, magnified views of selected regions. Due to locally weaker signals, blebs (red arrowheads) are poorly detected with the level set active contour algorithm.

The gradient vector flow algorithm performs better, but fails to detect subtler blebs (region 1) and does not accurately delineate deep crevices surrounding blebs (regions 2,3).

BleND was the only algorithm that could attain GT precision (Fig 5B). In general, the LS active contour algorithm resulted in inaccurate segmentations of nuclei with lower intensity or intensity gradients (Fig 5B).

Clustering of normal and dysmorphic nuclei

After validation of the algorithm, morpho-textural features were extracted from the segmented nuclei of HDF-NCP (Fig 6A). Hierarchical clustering of the normalized data based on Manhattan distance and Ward's clustering method, identified two major clusters, largely corresponding to the normal and dysmorphic nuclei, with ~ 89% correspondence to the manually assigned classes (Fig 6B). Visual inspection of incorrectly clustered nuclei revealed that their classification is often dubious, due to the presence of rough boundaries (in normal nuclei) or absence of overt blebs (for aberrant nuclei) (Fig 6C). The heatmap revealed a higher correlation of several shape parameters and the lack of correlation with textural features, as could be expected for this specific dataset since the HDF-NCP nuclei do not show any discriminating intensity-based characteristics. Indeed, EFD, curvature, solidity and circularity features comprised the most relevant information. This qualitative evaluation was confirmed by PCA, which revealed that the first principal component (PC1), explaining 29.25% of the variance within the dataset, contained no texture feature and determined the strongest direction of class separation. PC1 was defined by the EFD, curvature, solidity and circularity features (Fig 6D and 6E). The performance of hierarchical clustering could however not be improved significantly by using a reduced feature subset from PC1 (absolute correlation with PC1 larger than 0.7) (data not shown). As could be expected, EFD parameters dominated both clustering and PCA. Tracing the values of the summed EFD back to the segmented nuclei revealed a strong correlation between the severity of the shape alterations and the EFD value (Fig 7A). However, the summed EFD score by itself was not enough to distinguish all nuclei, since there was an overlap in the 0.4–0.5 range between normal and dysmorphic nuclei, thus calling for integration of other morphological parameters. In the NCP dataset, textural parameters had little impact on the classification result, but there were conditions were texture did significantly add to the discriminatory power. Indeed, in many HDF-NULL cells, nuclei are not only

Table 1. Performance errors of automatic segmentation methods.

Method	Class	Error
BleND*	Normal	0.935±0.222
	Dysmorphic	0.936±0.131
LS**	Normal	1.607±0.861
	Dysmorphic	2.678±2.243
GVF***	Normal	1.022±0.345
	Dysmorphic	1.153±0.386

*Threshold settings: Global = Triangle, Local = Mean

**LS = Level Set Active Contour: Advection 2.20, Curvature 0.10, Grayscale Tolerance 0.01, Convergence 0.0030

***GVF = Gradient Vector Flow Active Contour: Iterations 400, Tension or alpha 0, Rigidity or beta 0, External force or kappa, 30.
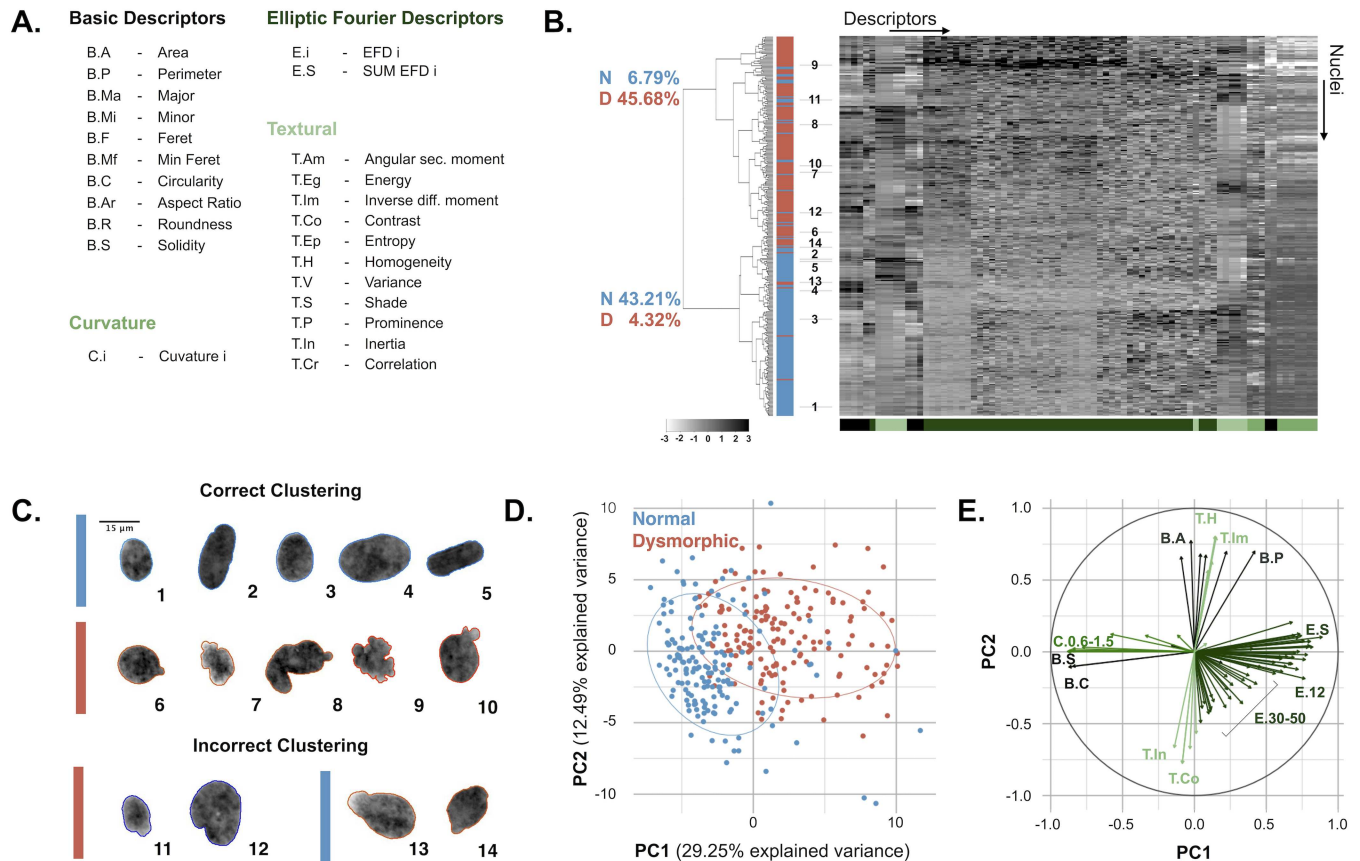
Fig 6. Unsupervised classification of automatically segmented nuclei. (A) Overview of the morpho-textural feature set that was extracted from 324 segmented nuclei; (B) Heatmap representing the grayscale-coded z-scores of all the features (columns) for all individual nuclei (rows). Hierarchical clustering on this dataset largely, but not completely, separates normal (blue) from dysmorphic (red) nuclei populations as indicated by the dendrogram on the left. (C) Example nuclei that have been correctly or incorrectly clustered. Colored outlines represent the manually assigned class, whereas the colored bar represents the assigned class by clustering (blue: normal and red: dysmorphic nuclei). Numbers link segmentations of selected nuclei to their position in the heatmap; (D) Principal component analysis of the data set yields two distinct but not fully separated clusters for the two classes as illustrated by a bi-plot explaining 42% of the variance. Discrimination of the two groups is predominantly in the direction of PC1; (E) The factor map reveals correlated features in PC space.

doi:10.1371/journal.pone.0170688.g006

dysmorphic, but also show local chromatin ruffling. This feature could effectively be picked up by textural features and allowed discrimination of morphologically similar nuclei (Fig 7B).

Supervised classification enables robust detection of dysmorphic nuclei

Using the selected feature set, a classifier was built for predicting nuclear dysmorphy. Various classification algorithms were assessed; their optimal MCR and the FNR are listed in Table 2. A support vector machine with a radial basis function (i.e. Gaussian) kernel yielded the best FNR, whereas a random forest classifier (300 trees, 5 features) had the best MCR on a training set through 10-fold cross-validation. On an independent test set, the support vector algorithm attained the best results with an MCR of 4.65% and a FNR of 0.92%. Because of their high performance, trainable SVM and RF classifiers (WEKA library [47]) were integrated in BleND. Herein, segmented nuclei can be assigned with a user-defined label through a graphical user interface. After manually categorizing a set of nuclei, a classifier is built and used to predict the classes for a larger set of segmented nuclei. Predictions can be improved by iterative addition of new manually assigned classes and classifier building (S1 Fig).
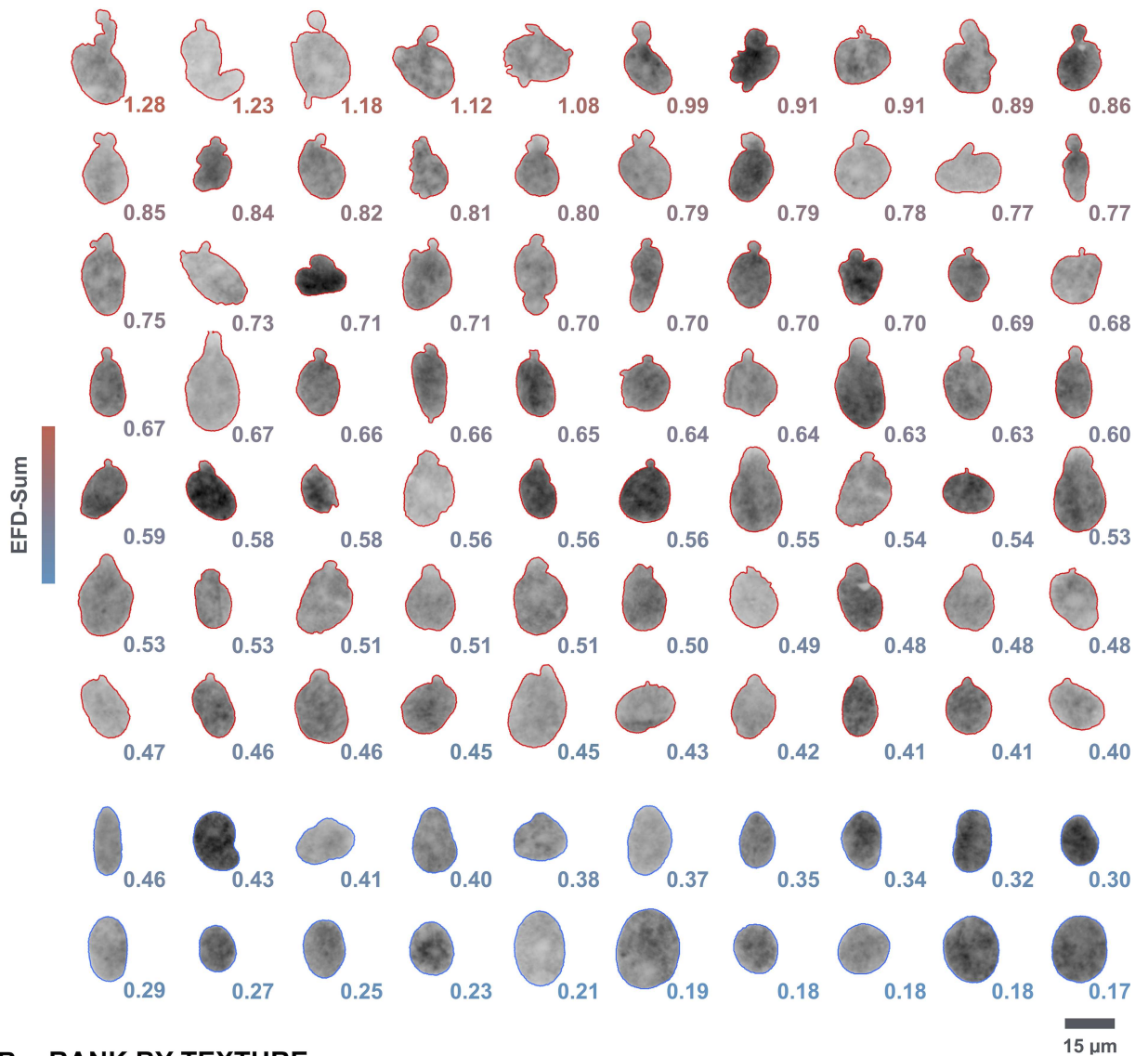
Fig 7. Discrimination of dysmorphic nuclei based on elliptic Fourier descriptors. (A) Morphology-based ranking of HDF-NCP nuclei. Both dysmorphic (red) and normal (blue) nuclei of HDF-NCP are ranked according to their summed EFD value (color coded). Severely deformed nuclei have higher EFD values than nuclei with small blebs, which in turn have larger EFD values than regular, ovoid-shaped nuclei. (B) Texture-based ranking for HDF-NULL cells. Dysmorphic nuclei are characterized by an intensity gradient due to an chromatin ruffling. Normal and aberrant nuclei of comparable shape (EFD value in italic and in brackets) can be distinguished based on the value of the entropy texture parameter (color coded).

doi:10.1371/journal.pone.0170688.g007

Table 2. Classification performance of different classifiers* on a training and test set using 10-fold cross-validation.

	LDA	QDA	RDA	MDA	FDA	NB	SVM	BAG	BOO	RF
TRAIN MCR	10.6	9.3	9.7	7.9	9.25	10.6	7.9	7.8	9.3	6.0
TRAIN FNR	7.9	0.9	7.0	4.1	4.2	5.1	0.9	2.3	4.6	1.9
TEST MCR	13.9	11.1	15.7	11.1	10.2	11.1	4.6	12.0	13.9	10.2
TEST FNR	10.2	4.6	13.0	7.4	5.6	6.5	0.9	5.6	7.4	4.6

*Settings: RDA: lambda 0.8997867, MDA: 3 subclasses, NB: Epanechnikov, SVM: Gaussian, BAG: 50 trees, BOO: 4561 iterations, RF: 500 trees, 3 features)

doi:10.1371/journal.pone.0170688.t002

Discussion

Dysmorphic nuclei are characteristic for a wide range of pathologies such as cancer, viral infections and nuclear envelopathies. Automated recognition and analysis of these nuclei may enhance the efficiency of cell-based microscopy experiments aimed at unraveling mechanisms underlying pathology. To this end, we wrote an algorithm that is tailored towards segmentation of dysmorphic nuclei and can be used for a wide variety of cell types acquired with different image modalities. Based on an integrated error score, we have shown that BleND attained a precision that matched the ground truth, when taking into account an inter-individual variability of 7%. The algorithm was further used to build a classifier that accurately predicts whether a nucleus is normal or dysmorphic.

Crevices and blebs that define dysmorphic nuclei entail major challenges due to their possible small size and lower intensity. Since there is no prior knowledge about the location, shape or intensity other algorithms described in literature are less suitable for this purpose. Segmentation algorithms relying on shape-based seed detection are not applicable to detect dysmorphic nuclei, since the shape of these nuclei strongly deviates from the normal convex shape [16,17]. Other algorithms use intensity information for the segmentation of the nuclei. However, local intensity minima in blebs can negatively influence the result of level sets-based methods as proven when comparing to BleND [48].

An algorithm that is optimised for the detection of nuclei with small aberrations has been described and is based on an GVF active contour algorithm [12]. As shown, error scores of this algorithm were similar to those of BleND for normal nuclei, but only BleND could attain the precision of manual delineations for dysmorphic nuclei. In line with the results of Driscoll *et al*, we found that curvature and solidity are good predictors of nuclear blebbing [12]. However, we now also show that EFD parameters are stronger correlates of nuclear dysmorphy and that the sum of these features correlates strongly with the severity of the deformation. In addition, BleND offers an alternative approach based on a simple DAPI staining rather than an immunofluorescence labelling of lamin A/C, making it more amenable for rapid, routine screening and multiplexing.

We have supported BleND with a framework for quantitative estimation of segmentation performance. The scoring system is based on scaled error metrics that describe the difference between the automated segmentation and user defined ground truths, and makes it possible to select the best threshold combination for the image data sets at hand. Our results demonstrate that a combination of global and local thresholding outperforms a single thresholding step. For contour refinement, we make use of a dynamic programming approach that is preceded by a straightening step and edge enhancement (derivative) of the initial contour. Other dynamic programming alternatives described in literature use polar transformation instead of bilinear interpolation for straightening of the edge [17,18]. This requires a centre and contour point as well as a mean radius to be defined in order to transform Cartesian into polar

coordinates. Since dysmorphic nuclei are characterized by their non-circularity, a polar transformation seems unfit. Normalisation of the columns equalizes the weight of all pixel values, causing the original lower intensities of edges in blebs or crevices to have an equal influence on the average path strength. In addition, the algorithm allows propagation angles greater than 45˚ to accurately describe crevices surrounding nuclear blebs. Segmentation results using the contour refinement algorithm were significantly better than those using only a thresholding step.

Morpho-textural features were extracted from the segmented nuclei and used for supervised classification with an accuracy up to 95%. Classification of the HDF-NCP nuclei was mainly determined by features that describe the shape of the nuclei such as curvature and the EFD descriptors. This is not surprising, as this type of cells does not show major textural alterations. However, other cell types such as HDF-NULL cells do show biologically relevant intensity variations (reflecting chromatin ruffling [21]) that may need to be discriminated. We showed that for similarly shaped nuclei, texture metrics such as entropy can discriminate chromatin ruffling. Thus, including these textural features makes BleND applicable to a broad range of cell types showing nuclear alterations.

The automated recognition is perfectly suited to be implemented in high-content perturbation screens that score nuclear shape changes associated with knockdown of specific genes [49,50] or treatment with chemical compounds [51]. A next logical step would be to integrate this automated recognition algorithm in an intelligent imaging workflow [52–54]. During live cell imaging, relevant events may be missed, since the observer manually has to define a region of interest before starting the experiment and the time resolution per well or spot is limited [55]. However, when the scope of the experiment can be limited to only those nuclei of interest (*in casu*, dysmorphic nuclei), the efficiency may be significantly increased. Feedback regulation between the microscope and the algorithm can result in automatic recognition of dysmorphic nuclei and subsequent initiation of an appropriate acquisition. A first step towards an integrated, broadly applicable intelligent imaging workflow, is the implementation of an iterative machine learning scheme in which a classifier can be trained on the fly, i.e. whilst images are being acquired [50]. As proof of principle, such an iterative learning process was integrated in BleND, resulting in classification scores similar to those of the independent data analysis that was performed. This workflow can be the starting point for high-resolution follow-up of more deformable nuclei [56], or even more complex imaging schemes such as selective, functional imaging (FRET, FCS, FRAP. . .) of dysmorphic nuclei, as has been demonstrated for mitotic phenotypes [57].

In conclusion, the proposed method can accelerate both fundamental research as well as diagnostics of the broad range of pathologies that are linked to nuclear dysmorphy.

Supporting Information

S1 Fig. Improved predictions by iterative training of implemented classifier. The misclassification rate (MCR) declines after iterative training of the implemented classification algorithm through a graphical user interface. In this example, a random forest classifier was used (100 trees, 7 features).
(TIFF)

Acknowledgments

Author Contributions

Conceptualization: MV HC JDV WP WDV.

Data curation: MV JR WDV.

Formal analysis: MV WDV.

Funding acquisition: WDV MV WP.

Investigation: MV JR HC.

Methodology: MV HC JDV JR WDV.

Project administration: WDV.

Resources: WDV WP.

Software: MV HC.

Supervision: WDV.

Validation: MV.

Visualization: MV WDV.

Writing – original draft: MV JDV WDV.

Writing – review & editing: MV WDV.

References

1. Zink D, Fischer AH, Nickerson JA. Nuclear structure in cancer cells. Nat Rev Cancer. 2004; 4: 677–87. doi: 10.1038/nrc1430 PMID: 15343274

2. Chow KH, Factor RE, Ullman KS. The nuclear envelope environment and its cancer connections. Nat Rev Cancer. 2012; 12: 196–209. doi: 10.1038/nrc3219 PMID: 22337151

3. Broers JLV, Ramaekers FCS, Bonne G, Ben Yaou R, Hutchison CJ. Nuclear lamins: Laminopathies and their role in premature ageing. Physiol Rev. 2006; 86: 967–1008. doi: 10.1152/physrev.00047.2005 PMID: 16816143

4. Burke B, Stewart CL. Life at the edge: the nuclear envelope and human disease. Nat Rev Mol Cell Biol. 2002; 3: 575–85. doi: 10.1038/nrm879 PMID: 12154369

5. De Vos WH, Houben F, Kamps M, Malhas A, Verheyen F, Cox J, et al. Repetitive disruptions of the nuclear envelope invoke temporary loss of cellular compartmentalization in laminopathies. Hum Mol Genet. 2011; 20: 4175–86. doi: 10.1093/hmg/ddr344 PMID: 21831885

6. Houben F, De Vos WH, Krapels IPC, Coorens M, Kierkels GJJ, Kamps MAF, et al. Cytoplasmic localization of PML particles in laminopathies. Histochem Cell Biol. 2013; 139: 119–34. doi: 10.1007/s00418-012-1005-5 PMID: 22918509

7. Denais CM, Gilbert RM, Isermann P, McGregor AL, Lindert te M, Weigelin B, et al. Nuclear envelope rupture and repair during cancer cell migration. Science. 2016; 352: 353–8. doi: 10.1126/science.aad7297 PMID: 27013428

8. Robijns J, Molenberghs F, Sieprath T, Corne TDJ, Verschuuren M, De Vos WH. In silico synchronization reveals regulators of nuclear ruptures in lamin A/C deficient model cells. Sci Rep. 2016; 6: 1–11.

9. Hatch E, Hetzer M. Breaching the nuclear envelope in development and disease. J Cell Biol. 2014; 205: 133–41. doi: 10.1083/jcb.201402003 PMID: 24751535

10. de Noronha CM, Sherman MP, Lin HW, Cavrois MV, Moir RD, Goldman RD, et al. Dynamic disruptions in nuclear envelope architecture and integrity induced by HIV-1 Vpr. Science. 2001; 294: 1105–8. doi: 10.1126/science.1063957 PMID: 11691994

11. De Vos WH, Van Neste L, Dieriks B, Joss GH, Van Oostveldt P. High content image cytometry in the context of subnuclear organization. Cytometry A. 2010; 77: 64–75. doi: 10.1002/cyto.a.20807 PMID: 19821512

12. Driscoll MK, Albanese JL, Xiong Z-M, Mailman M, Losert W, Cao K. Automated image analysis of nuclear shape: what can we learn from a prematurely aged cell? Aging (Albany NY). 2012; 4: 119–32.

13. Xu C, Prince JL. Snakes, shapes, and gradient vector flow. IEEE Trans Image Process. 1998; 7: 359–69. doi: 10.1109/83.661186 PMID: 18276256

14. Lesko M, Kato Z, Nagy A, Gombos I, Torok Z, Vigh L Jr, et al. Live Cell Segmentation in Fluorescence Microscopy via Graph Cut. ICPR. 2010; 1485–8.

15. Boykov Y, Funka-Lea G. Graph Cuts and Efficient N-D Image Segmentation. Int J Comput Vis. 2006; 70: 109–31.

16. Arslan S, Ersahin T, Cetin-Atalay R, Gunduz-Demir C. Attributed Relational Graphs for Cell Nucleus Segmentation in Fluorescence Microscopy Images. IEEE Trans Med Imaging. 2013; 32: 1121–31. doi: 10.1109/TMI.2013.2255309 PMID: 23549886

17. Nandy K, Gudla PR, Lockett SJ. Automatic segmentation of cell nuclei in 2D using dynamic programming. Proceedings of 2nd Workshop on Microsopic Image Analysis with Applications in Biology. 2007; 6718–21.

18. Baggett D, Nakaya MA, McAuliffe M, Yamaguchi TP, Lockett S. Whole cell segmentation in solid tissue sections. Cytometry A. 2005; 67A: 137–43.

19. Verstraeten VLRM, Broers JLV, van Steensel MAM, Zinn-Justin S, Ramaekers FCS, Steijlen PM, et al. Compound heterozygosity for mutations in LMNA causes a progeria syndrome without prelamin A accumulation. Hum Mol Genet. 2006; 15: 2509–22. doi: 10.1093/hmg/ddl172 PMID: 16825282

20. Muchir A, van Engelen BG, Lammens M, Mislow JM, McNally E, Schwartz K, et al. Nuclear envelope alterations in fibroblasts from LGMD1B patients carrying nonsense Y259X heterozygous or homozygous mutation in lamin A/C gene. Exp Cell Res. 2003; 291: 352–62. PMID: 14644157

21. De Vos WH, Houben F, Hoebe RA, Hennekam R, van Engelen B, Manders EMM, et al. Increased plasticity of the nuclear envelope and hypermobility of telomeres due to the loss of A-type lamins. Biochim Biophys Acta. 2010; 1800: 448–58. doi: 10.1016/j.bbagen.2010.01.002 PMID: 20079404

22. Cornelissen F, Verstraelen P, Verbeke T, Pintelon I, Timmermans J-P, Nuydens R, et al. Quantitation of chronic and acute treatment effects on neuronal network activity using image and signal analysis: toward a high-content assay. J Biomol Screen. 2013; 18: 807–19. doi: 10.1177/1087057113486518 PMID: 23606652

23. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012; 9: 676–82. doi: 10.1038/nmeth.2019 PMID: 22743772

24. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012; 9: 671–5. PMID: 22930834

25. Huang LK, Wang M. Image Thresholding by Minimizing the Measures of Fuzziness. Pattern Recognit. 1995; 28: 41–51.

26. Prewitt JM, Mendelsohn ML. The analysis of cell images. Ann N Y Acad Sci. 1966; 128: 1035–53. PMID: 5220765

27. Ridler TW, Calvard S. Picture Thresholding Using an Iterative Selection Method. IEEE Trans Syst Man Cybern. 1978; 8: 630–2.

28. Li CH, Tam P. An iterative algorithm for minimum cross entropy thresholding. Pattern Recognit Lett. 1998; 19: 771–6.

29. Kapur JN, Sahoo PK, Wong A. A New Method for Gray-Level Picture Thresholding Using the Entropy of the Histogram. Computer Vision Graphics and Image Processing. 1985; 29: 273–85.

30. Glasbey CA. An Analysis of Histogram-Based Thresholding Algorithms. CVGIP—Graph Model Im. 1993; 55: 532–7.

31. Kittler J, Illingworth J. Minimum Error Thresholding. Pattern Recognit. 1986; 19: 41–7.

32. Tsai WH. Moment-Preserving Thresholding—a New Approach. Comput Vis Graph Image Process. 1985; 29: 377–93.

33. Otsu N. A Threshold Selection Method from Gray-Level Histograms. Automatica. 1975; 11: 23–7.

34. Doyle W. Operations Useful for Similarity-Invariant Pattern Recognition. J Acm. 1962; 9: 259–267.

35. Shanbhag AG. Utilization of Information Measure as a Means of Image Thresholding. CVGIP—Graph Model Im. 1994; 56: 414–9.

36. Zack GW, Rogers WE, Latt SA. Automatic-Measurement of Sister Chromatid Exchange Frequency. J Histochem Cytochem. 1977; 25: 741–53. PMID: 70454

37. Yen JC, Chang FJ, Chang SA. A New Criterion for Automatic Multilevel Thresholding. IEEE Trans Image Process. 1995; 4: 370–8. doi: 10.1109/83.366472 PMID: 18289986

38. Kocsis E, Trus BL, Steer CJ, Bisher ME, Steven AC. Image Averaging of Flexible Fibrous Macromole-cules—the Clathrin Triskelion Has an Elastic Proximal Segment. J Struct Biol. 1991; 107: 6–14. PMID: 1817611

39. Vincent L, Soille P. Watersheds in Digital Spaces—an Efficient Algorithm Based on Immersion Simula-tions. IEEE Trans Pattern Anal Mach Intell. 1991; 13: 583–98.

40. Dubuisson M-P, Anil JK. A modified Hausdorff distance for object matching. Conference: Pattern Rec-ognition, Conference A: Computer Vision & Image Processing., Proceedings of the 12th IAPR Interna-tional Conference on, Volume: 1. 1994.

41. Dice LR. Measures of the Amount of Ecologic Association Between Species. Ecology. 1945; 26: 297–302.

42. Furuta N, Ninomiya S, Takahashi N, Ohmori H, Ukai Y. Quantitative-Evaluation of Soybean (Glycine-Max L Merr) Leaflet Shape by Principal Component Scores Based on Elliptic Fourier Descriptor. Breed Sci. 1995; 45: 315–20.

43. Haralick RM, Karthikeyan S, Dinstein IH. Textural features for image classification. IEEE Trans Syst Man Cybern. 1973; 3: 610–21.

44. R Development Core Team. R: A language and environment for statistical computing. 2008. http://www.R-project.org

45. Choi S, Wang W, Ribeiro AJS, Kalinowski A, Gregg SQ, Opresko PL, et al. Computational image analy-sis of nuclear morphology associated with various nuclear-specific aging disorders. Nucleus. 2014; 2: 570–9.

46. Kass M, Witkin A, Terzopoulos D. Snakes—Active Contour Models. Int J Comput Vis. 1987; 1: 321–31.

47. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The WEKA data mining software: an update. ACM SIGKDD explorations newsletter. 2009; 11: 10–8.

48. Gudla PR, Nandy K, Collins J, Meaburn KJ, Misteli T, Lockett SJ. A high-throughput system for seg-menting nuclei using multiscale techniques. Cytometry A. 2008; 73: 451–66. doi: 10.1002/cyto.a.20550 PMID: 18338778

49. Neumann B, Walter T, Heriche JK, Bulkescher J, Erfle H, Conrad C, et al. Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. Nature. 2010; 464: 721–7. doi: 10.1038/nature08869 PMID: 20360735

50. Jones TR, Carpenter AE, Lamprecht MR, Moffat J, Silver SJ, Grenier JK, et al. Scoring diverse cellular morphologies in image-based screens with iterative feedback and machine learning. Proc Natl Acad Sci USA. 2009; 106: 1826–31. doi: 10.1073/pnas.0808843106 PMID: 19188593

51. Kubben N, Brimacombe KR, Donegan M, Li Z, Misteli T. A high-content imaging-based screening pipe-line for the systematic identification of anti-progeroid compounds. Methods. 2015; 96: 46–58. doi: 10.1016/j.ymeth.2015.08.024 PMID: 26341717

52. Wenus J, Duessmann H, Paul P, Kalamatianos D, Rehm M, Wellstead PE, et al. ALISSA: an automated live-cell imaging system for signal transduction analyses. BioTechniques. 2009; 47: 1033–40. doi: 10.2144/000113247 PMID: 20041856

53. Conrad C, Wuensche A, Tan TH, Bulkescher J, Sieckmann F, Verissimo F, et al. Micropilot: automation of fluorescence microscopy-based imaging for systems biology. Nat Methods. 2011; 8: 246–89. doi: 10.1038/nmeth.1558 PMID: 21258339

54. Jackson C, Glory-Afshar E, Murphy RF, Kovacevic J. Model building and intelligent acquisition with application to protein subcellular location classification. Bioinformatics. 2011; 27: 1854–9. doi: 10.1093/bioinformatics/btr286 PMID: 21558154

55. Pepperkok R, Ellenberg J. Innovation—High-throughput fluorescence microscopy for systems biology. Nat Rev Mol Cell Biol. 2006; 7: 690–6. doi: 10.1038/nrm1979 PMID: 16850035

56. De Vylder J, De Vos WH, Manders EM, Philips W. 2D mapping of strongly deformable cell nuclei-based on contour matching. Cytometry A. 2011; 79A: 580–8.

57. Rohde GK, Ribeiro AJS, Dahl KN, Murphy RF. Deformation-based nuclear morphometry: capturing nuclear shape variation in HeLa cells. Cytometry A. 2008; 73: 341–50. doi: 10.1002/cyto.a.20506 PMID: 18163487