

SEGMENTATION OF CELL STRUCTURES USING MODEL-BASED SET COVERING WITH ITERATIVE REWEIGHTING

Peter Markowsky¹ Svenja Reith² Tabea E. Zuber¹ Rainer König³ Karl Rohr² Christoph Schnörr¹

¹Image and Pattern Analysis Group, Heidelberg University, Germany

²Biomedical Computer Vision Group, Dept. Bioinformatics and Functional Genomics, University of Heidelberg, BIOQUANT, IPMB, and DKFZ Heidelberg, Germany

³Center for Sepsis Control and Care (CSCC), Jena University Hospital, Germany

ABSTRACT

We present a new method for cell segmentation which combines a marked point process model with a combinatorics-based method of finding global optima. The method employs an energy term that assesses possible segmentations by their fidelity to both local image information and a simple model of cell interaction, and we use a randomized iterative reweighting technique for its minimization. Our approach was successfully applied to cell microscopy images of varying difficulty and experimentally compared with both a standard segmentation method as well as a method based on Multiple Birth and Cut [1]. The proposed method is found to improve upon previous approaches.

Index Terms— Cell Segmentation, Iterative Candidate Reweighting, Randomized Set Covering, Marked Point Process

1. INTRODUCTION

The task of segmenting cell nuclei or whole cells in microscopy images poses a major challenge for current biomedical imaging applications. As a result of new techniques for microscopy imaging such as SPIM [2], the amount of image data has been growing continuously over the last years. Applications such as high-throughput screenings have made manual segmentation increasingly infeasible, giving rise to the development of various automated approaches. In the context of cell (nuclei) segmentation, especially densely populated microscopy images with a large number of clustering cells are challenging for current segmentation algorithms.

Previous approaches for cell segmentation are primarily based on intensity thresholding, feature detection, morphological filtering, region accumulation or deformable model fitting (see [3] for an overview). Some recent efforts focus on the generation of seeds required to initialize region accumulation methods such as marker-controlled watershed as well as deformable model fitting approaches such as active contours. The segmentation results of these methods strongly depend on a proper choice of seed points. One strategy for seed generation entails calculating seeds directly from the image, for example, based on the image gradient for initializing active contours.

In cases where local image information is not sufficient, an a priori model of object interaction, for example, based on Marked Point Processes (MPP), can be incorporated into the segmentation. Among others, [4],[5] employed this Bayesian approach: A good segmentation is defined as the set of objects – or “candidates” – that minimize an energy function. This energy function measures fidelity to both image data and a simplified model of pairwise interaction

of the segmented objects. Resulting energies are, however, usually highly non-convex, and global optimization is not computationally feasible in practice. A recent approach [6],[1] solves this problem by iteratively selecting a random set of candidates, and performing the minimization locally on this set. However, sampling globally optimal candidates from a uniform distribution over a large number of candidates requires many resampling steps. In [6] this problem is addressed by alternating between sampling completely random candidates and local perturbations of the previous segmentation.

In this contribution, we present Model-Based Set Covering (MBSC), a different approach compared to [6] for achieving faster convergence: A model-based energy is iteratively minimized over a random set of candidates, but the distribution according to which they are sampled is reweighted in each step. This update is based on the Geometric Set Cover algorithm in [7]. The idea is to regard segmentation as the problem of finding a minimal set of candidates that cover all object pixels. The probability of sampling such a set can be shown to systematically increase in each step under mild conditions, leading to convergence in almost surely finite time. The Geometric Set Cover algorithm can be seen as a special case of the multiplicative weights method described in [8].

In the following sections 2 and 3, we briefly discuss our choice of candidate model for segmentation and the energy model. Then, in section 4, we introduce our new method for updating the sampling probabilities. Section 5 provides a quantitative evaluation using fluorescently labeled neuroblastoma cell images of different cell density as well as a comparison with previous approaches.

2. SETTING AND CHOICE OF CANDIDATES

In the following, a “candidate” is a set of points that represent possible segmentations of single cells. Given a set of discrete candidates \mathcal{C} , our goal is to find those that fit the image best without violating our a priori model of cell interaction.

For simplicity, as well as computational speed, we confine ourselves to the easily parameterizable case of circles with varying centers and radii as candidates for cell shapes. However, the use of either shapes of higher parameter dimension, for example, ellipses, or a set of candidates that is not parameterized at all, as used, for example, in [9], is costly only in computational time, and the calculation of pairwise overlap is slightly more complicated, without presenting any theoretical obstruction.

3. INTERACTION ENERGY MODEL

Given a finite subset of candidates $X \subseteq \mathcal{C}$, our energy function is composed of a unary energy term U_1 as well as a pairwise interaction energy U_2 :

$$U(X) := \underbrace{\sum_{c_i \in X} D(c_i)}_{U_1(X)} + \underbrace{\sum_{\substack{c_i, c_j \in X \\ c_i \neq c_j}} \chi_{[0, \infty)}(d(c_i, c_j))}_{U_2(X)} \quad (1)$$

where d measures the distance between two candidates c_i and c_j , and

$$\chi_{[0, \infty)}(r) := \begin{cases} 0 & \text{if } r \geq 0 \\ \infty & \text{else} \end{cases} \quad (2)$$

is the characteristic function. U_1 assigns a function D rewarding fidelity to image data to each possible candidate, while U_2 excludes pairs of candidates that interact in a way we deem unrealistic with an infinite penalty. Specifically, if p_i and r_i are defined as center point and radius of a candidate circle c_i , we define

$$d(c_i, c_j) := \|p_i - p_j\| - t(r_i + r_j) \quad (3)$$

as a measure of pairwise overlap relative to circle radius. The parameter t governs the amount of overlap that is still acceptable, $t = 1$ meaning no overlap at all. Our choice of the data fidelity term D is dependent on the sampling probabilities output by the Set Cover update and will be explained in detail in section 4.

In earlier works (for example, [6]) graph cut methods were used for energy minimization. Application of this method is, however, limited to minimizing over two subsets X_1 and X_2 , each of which must internally be non-interacting, i.e. $U_2(X_1) = U_2(X_2) = 0$. Using a binary global optimizer allows for more flexibility in the number and choice of candidates over which the energy is minimized. It further eliminates the need to calculate a set of non-interacting candidates as a preprocessing step to energy minimization. Tree Reweighted Message Passing (TRWS) is our binary optimizer of choice, primarily because its exceptional speed is convenient for repeated minimization (see [10] for a detailed description of the method).

Our approach consists of iteratively sampling a finite set of candidates according to a probability distribution μ_n , finding a subset $\mathcal{C}_{(n)}^*$ that minimizes the energy over both the newly sampled candidates and the optimum calculated in the previous step, and updating the probability distribution μ_n using the update described in the next section. See Algorithm 1 for an overview.

Algorithm 1: Model-Based Set Covering

```

Sample  $\mathcal{C} \supset \mathcal{C}_{(0)} \sim \mu_0$ 
while  $n < n_{\max}$  do
  Sample a new set of candidates  $N_{(n)} \sim \mu_n$ 
  Combine with old optimum:  $\mathcal{C}_{(n)} = N_{(n)} \cup \mathcal{C}_{(n-1)}^*$ 
  Minimize energy:  $\mathcal{C}_{(n)}^* = \operatorname{argmin}_{X \subseteq \mathcal{C}_{(n)}} U(X)$ 
  Update  $\mu_n$  using Algorithm 2
return  $\mathcal{C}_{(n_{\max})}^*$ 

```

4. PROBABILITY UPDATE USING RANDOMIZED SET COVERING

As mentioned before, the basic idea of the Set Cover update is to view the labeling of candidates as the problem of finding a minimal set of candidates that cover all object pixels. This minimal set cover is approximated by iteratively increasing the probability of covering those pixels which are covered with a low probability in the current step: Given $p \in P$, where P is the set of all object pixels – obtained, for example, by thresholding – we can define

$$\mathcal{C}^p := \{c \in \mathcal{C} : c \cap \{p\} \neq \emptyset\} \quad (4)$$

as the set of all candidates that potentially cover the pixel p . If p is not covered in a certain step, and the a priori probability of covering it is smaller than a predefined value ε , i.e.

$$\mu_n(\mathcal{C}^p) := \sum_{c \in \mathcal{C}^p} \mu_n(c) < \varepsilon \quad (5)$$

we double the current probabilities $\mu_n(c)$ of all candidates in this set, followed by normalization to obtain a new distribution $\mu_{(n+1)}$. The set cover probability update is summarized in Algorithm 2.

Algorithm 2: Set Cover Probability Update

```

set  $\bar{P}_c = P \setminus \bigcup \mathcal{C}_n$ 
if  $\bar{P}_c \neq \emptyset$  then
  pick any point  $p \in \bar{P}_c$ 
  if  $\mu_n(\mathcal{C}^p) < \varepsilon$  then
     $\omega(c) \leftarrow 2\mu_{(n+1)}(c), \forall c \in \mathcal{C}^p$ 
     $\omega(c) \leftarrow \mu_{(n+1)}(c), \forall c \notin \mathcal{C}^p$ 
    Normalize:  $\mu_{(n+1)}(c) \leftarrow \frac{\omega(c)}{\sum_{c'} \omega(c')}, \forall c \in \mathcal{C}$ 
  else
     $\mu_{(n+1)} \leftarrow \mu_{(n)}$ 

```

If a set \mathcal{O} of k optimal candidates exists and $\varepsilon < 2^{\frac{1}{k}} - 1$, condition (5) ensures that $\mu_n(\mathcal{O})$ grows in each step. Convergence is guaranteed in almost surely finite time if the probability of condition (5) being satisfied is not zero. For details on the proof of convergence, see [7]. Note that while the original setting does not include object interaction, the proof of convergence only relies on the existence of an optimal subset, and the probabilities of its members being updated in each step, and thus still holds. Similarly, the update can be performed for multiple points p_i in one step, as long as the sets \mathcal{C}^{p_i} do not intersect.

In practice, a non-zero probability for condition (5) can be easily verified in a stepwise manner. The choice of a value of ε that guarantees convergence requires estimating an upper bound of the cardinality of a set of optimal candidates, i.e. the number of cells. This bound does not need to be precise, however. For our application, we estimate the number of cells as a fixed multiple of the number of connected components of the thresholded cell image (either 1.4 or 2.4 was used as a factor depending on cell density).

4.1. Data Term

The Set Cover algorithm is constructed to assign large probabilities to members of a minimal covering set, i.e. candidates that cover a lot of object pixels that others do not. This also means that the output probabilities can serve as an implicitly constructed data term D , see figure 1 for an example.

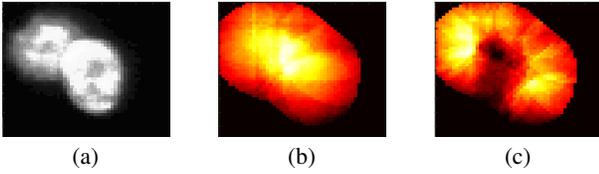


Fig. 1: Small section of a cell image with two touching cells. (a) Original image. (b) Probabilities assigned to possible centers of larger candidates and (c) smaller ones. For large candidates, one circle placed in the center of both cells is a minimal covering set, and the algorithm assigns large probabilities to this region. For those candidates of approximately the cell size, a minimal covering set consists of two candidates placed roughly in the center of each cell.

For our application, we choose a rescaled version of the probabilities μ_n as the data term. Specifically, given a set of sampled candidates X , we take the logarithm of μ_n , set the overall minimum to 0, the maximum to 1, and take the negative to obtain low values of the data term for candidates with high probabilities:

$$D'(c) := \log(\mu_n(c)) - \min_{c \in X} (\log \mu_n(c)) \quad (6)$$

$$D(c) = - \frac{D'(c)}{\max_{c \in X} D'(c)} \quad (7)$$

Rescaling avoids rounding errors both internally and in the TRWS minimization, as the increase of probabilities μ_n depends exponentially on the number of updates.

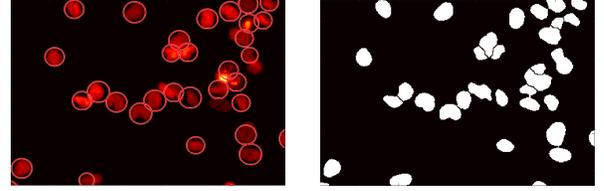
Constructing the unary energy in such a way eliminates the need to separately calculate a data term for each candidate, which may lead to a significant speed-up, especially for a large number of candidates and a computational expensive data term. However, the main advantage is that for an optimal set of candidates \mathcal{O} , its probability $\mu_n(\mathcal{O})$ increases with each update, leading to a steadily decreasing data term. If the interaction energy is not chosen unrealistically, i.e. assigning an infinite penalty to pairs of optimal candidates, this means that as the number of probability updates increases, better candidates will not only be preferably randomly sampled, but also be more and more likely to be part of a stepwise energy minimum.

The main disadvantage of this method is that candidates that cover more points have a high chance of being updated more often, and thus be assigned a higher probabilities than their subsets, even if, for example, the ratio of covered object pixels to the total area of the candidate is bad. We address this problem by assigning a penalty depending on the radius of the candidate at the start of the algorithm, ensuring that if a smaller candidate covers the same object points as a larger one, it will be always preferably chosen.

5. EXPERIMENTAL RESULTS

5.1. Setup and Methodology

Recall that the set of candidates \mathcal{C} consists of circles defined by their points of origin $p_i \in P'$ and radii $r_i \in \mathcal{R}$. The set P' can naively be chosen as the set of all image pixels. We apply an isotropic Gaussian filter to the original image and subsequently use conservative thresholding on the result to eliminate all very unlikely candidate centers as a pre-processing step for all algorithms. We choose the 9 possible radii $\mathcal{R} = \{7, 8, \dots, 15\}$ based on an inspection of typical cell sizes. As a post-processing step, we perform marker-controlled



(a) Approximate segmentation using circular candidates (b) Segmentation result after seeded watershed transform

Fig. 2: Seeded watershed transform for postprocessing (medium data)

watershed segmentation on the thresholded image, using the centres of segmented candidates as seeds, to improve the local matching of candidate shape to image data. Since seeded watershed is prone to oversegmentation, we apply the algorithm locally on the connected components of the thresholded image, and ignore connected components that are smaller than 5 pixels (see figure 2 for an example). As mentioned in section 4, we estimate the number k of cells as a multiple of the connected components of the thresholded cell image: 1.4 for easy and medium images, and 2.4 for difficult ones. Furthermore, we heuristically set the parameters to $t=0.75$ (see (3)), $|\mathcal{C}_{(n)}| = 3k$ (see Algorithm 1), and the starting penalty for candidates of radius r to $q(r) = 4^{-r^2}$ (see the last paragraph of subsection 4.1).

We evaluate three different approaches: First, our proposed approach, referred to as MBSC; its starting probabilities are proportional to the output of the filter-based pre-processing multiplied by $q(r)$. Second, a version of the approach that uses static uniform sampling probabilities and the static data term for circles proposed in [9] (MBC) with parameter $d_0 = 0.1$, and third, an approach that uses the data term of MBC multiplied by $q(r)$ as starting probabilities, but is otherwise identical to the first approach (denoted as MBSC/MBC). Note that MBC is similar to the Multiple Birth and Cut algorithm proposed in [1], with the difference of using a slightly different data term, and an interaction term that defines overlap based on distance instead of area. Also, in contrast to [1], the repeated energy minimization is performed on a random set of size $3k$ using the TRWS minimizer, instead of employing graph cuts on two sets of non-interacting candidates of roughly size k , as used for all three approaches evaluated here. In addition, we compare the three approaches with a standard segmentation method, namely Otsu thresholding [11] as implemented in Fiji/ImageJ [12] with additional pre- and postprocessing (contrast enhancement, anisotropic filtering, and morphological opening to fill holes).

5.2. Data and Performance Measures

We apply the approaches to microscopy images of fluorescently labeled neuroblastoma cell lines (see [13] for a detailed description of the data). The images have a size of 1344 x 1024 pixels. For the evaluation, the image data is categorized into different degrees of difficulty: the images of the cell line SH-EP exhibit a small to medium cell count and little clustering (low to medium difficulty). The images of the cell line SK-N-BE(2)-C have a high cell count as well as many clustering cells and constitute more challenging data. The evaluation is performed on 18 images comprising 6 images of each category of difficulty, i.e. easy, medium, and difficult. An example of an image of category easy is shown in figure 3a.

The performance of the approaches is evaluated with respect to manual segmentation by a human expert. The values of the con-

	easy				medium				difficult			
	Otsu	MBC	MBSC/ MBC	MBSC	Otsu	MBC	MBSC/ MBC	MBSC	Otsu	MBC	MBSC/ MBC	MBSC
Recall	87.9	91.4	87.8	88.2	90.3	92.4	89.3	89.5	69.8	79.8	69.5	71
Precision	98.8	77.9	95.7	93.4	98.9	83.5	96.9	96.8	94.3	72.5	93.7	86.9
FNR	12.1	8.63	12.6	12.2	9.92	7.61	11	10.7	33.5	20.2	34	31
Dice	0.846	0.859	0.876	0.876	0.865	0.871	0.888	0.887	0.81	0.8	0.857	0.826
σ_{Dice}	± 0.014	± 0.007	± 0.005	± 0.004	± 0.016	± 0.012	± 0.01	± 0.009	± 0.017	± 0.025	± 0.008	± 0.047
Hausdorff	1.77	1.75	1.87	1.86	9.63	9.16	9.52	9.42	11.8	12	15.6	14.4
$\sigma_{\text{Hausdorff}}$	± 0.311	± 0.468	± 0.374	± 0.373	± 1.484	± 1.263	± 1.16	± 1.194	± 1.053	± 1.394	± 1.723	± 0.571
ARI	0.839	0.852	0.869	0.869	0.82	0.825	0.846	0.845	0.72	0.707	0.778	0.739
σ_{ARI}	± 0.014	± 0.008	± 0.005	± 0.004	± 0.019	± 0.014	± 0.011	± 0.011	± 0.033	± 0.042	± 0.021	± 0.071
MCN	74 \pm 14				419 \pm 50				641 \pm 45			

Table 1: Averaged values of different performance measures chosen for the segmentation approaches applied to "easy", "medium" and "difficult" cell images. FNR is the false negative rate, and MCN is the mean cell number per image.

fusion matrix (True Positives, False Positives, True Negatives, and False Negatives) are used to determine recall and precision values. Additionally, we investigate the performance of the approaches using a selection of measures described in [14]. We used an overlap-based metric (Dice coefficient, F-measure), a distance-based metric (Hausdorff distance), and a pair-counting-based metric (Adjusted Rand Index, ARI) to assess different characteristics of the segmentation results. For the Hausdorff distance we calculate the modified Hausdorff distance as proposed in [15] for the True Positives. All measures are normalized with respect to the number of cells or pixels in the manually segmented images and averaged over all images of the respective category of difficulty.

5.3. Results

On average, manual segmentation resulted in a total of 74 cells per image for "easy", 419 cells for "medium", and 641 cells for "difficult" data. The results for the different performance measures averaged over all images of the respective category of difficulty are summarized in table 1. For all approaches the results for difficult data are worse compared to easy and medium data. Although the MBC algorithm achieves the highest recall and lowest FNR, this is mainly due to a consistent oversegmentation as evident in the relatively low precision. In terms of precision, Otsu thresholding yields the best results, and both MBSC and MBSC/MBC perform better than MBC. Also, MBSC/MBC is better than MBSC. Except for difficult images, MBSC and MBSC/MBC perform comparably well. Concerning the Hausdorff distance, MBC performs better than both MBSC and MBSC/MBC, and it is outperformed by Otsu thresholding for difficult images. However, the results for the Hausdorff distance should be interpreted with caution due to the relatively large standard deviations. Both MBSC and MBSC/MBC consistently outperform Otsu thresholding as well as MBC with regard to the Dice coefficient and ARI. The improvement is more prominent for difficult images with a high cell density and a higher amount of clustering. MBSC/MBC yields similar or slightly better results compared to MBSC in terms of precision, Dice coefficient, and ARI. The results show that although better-tuned starting probabilities of MBSC/MBC do have a slight positive influence, the Set Cover probability update of MBSC and MBSC/MBC has a much stronger impact on the result, and the result generally does not depend on the initialization.

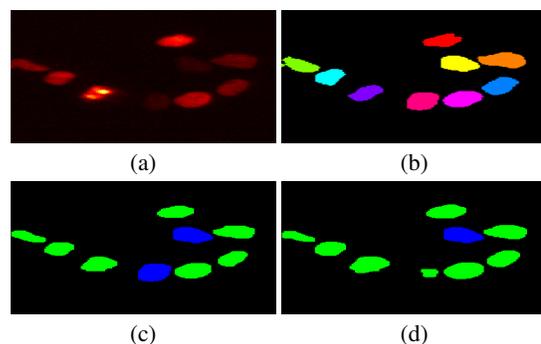


Fig. 3: (a) Section of an original cell microscopy image (easy data); (b) manual segmentation result; and segmentation result using (c) Otsu thresholding and (d) MBSC. For (c) and (d), cells marked in green are True Positives (TP), and those in blue are False Negatives (FN, taken from manual segmentation).

6. SUMMARY AND DISCUSSION

We have introduced an approach for cell segmentation which combines marked point process-based segmentation through repeated energy minimization with iterative candidate reweighting by a randomized set covering algorithm. The approach has several advantages: On the theoretical side, a stronger convergence is guaranteed, and it is not necessary to explicitly calculate a fitting data term. Experimentally, we found that our approach generally leads to improved results compared to a method based on the same energy model where the sampling probabilities and the data term are not updated. By combining our method with the data term previously presented in [9] the result was further improved. An extension of our approach would be the use of candidate shapes with a higher number of parameters, as well as a more sophisticated model of object interaction.

7. ACKNOWLEDGEMENTS

We thank Jonas Weidenhausen for providing the manual segmentation and the result of Otsu thresholding. We thank Nathalie Harder for selecting the images for the evaluation. This work has been supported by the RTG 1653 of the German Research Foundation (DFG), the AniS project of the German Federal Ministry of Education and Research (BMBF), and the BMBF project RNA-Code (e:Bio).

8. REFERENCES

- [1] E. Poulain, S. Prigent, E. Soubies, and X. Descombes, "Cells detection using segmentation competition," in *IEEE 12th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2015, pp. 1208–1211.
- [2] J. Huisken, J. Swoger, F. Del Bene, J. Wittbrodt, and E. H. K. Stelzer, "Optical sectioning deep inside live embryos by selective plane illumination microscopy," *Science*, vol. 305, no. 5686, pp. 1007–1009, 2004.
- [3] E. Meijering, "Cell segmentation: 50 years down the road [life sciences]," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 140–145, Sept 2012.
- [4] G. Dong and S. T. Acton, "Detection of rolling leukocytes by marked point processes," *Journal of Electronic Imaging*, vol. 16, no. 3, pp. 033013–033013(11), 2007.
- [5] A. Khan, S. Gould, and M. Salzmann, "A linear chain markov model for detection and localization of cells in early stage embryo development," in *Proc. IEEE Winter Conference on Applications of Computer Vision 2015*. IEEE, 2015, pp. 526–533.
- [6] E. Soubies, P. Weiss, and X. Descombes, "Graph cut based segmentation of predefined shapes: Applications to biological imaging," in *Pattern Recognition Applications and Methods*, pp. 153–170. Springer, 2015.
- [7] S. Har-Peled, *Geometric Approximation Algorithms*, AMS, 2011.
- [8] S. Arora, E. Hazan, and S. Kale, "The multiplicative weights update method: a meta-algorithm and applications.," *Theory of Computing*, vol. 8, no. 1, pp. 121–164, 2012.
- [9] X. Descombes, "Multiple objects detection in biological images using a marked point process framework," *Methods*, in press.
- [10] V. Kolmogorov, "Convergent tree-reweighted message passing for energy minimization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 10, pp. 1568–1583, 2006.
- [11] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, Jan 1979.
- [12] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, et al., "Fiji: an open-source platform for biological-image analysis," *Nature Methods*, vol. 9, no. 7, pp. 676–682, 2012.
- [13] R. Batra, N. Harder, S. Gogolin, N. Diessl, Z. Soons, C. Jäger-Schmidt, C. Lawrenz, R. Eils, K. Rohr, F. Westermann, et al., "Time-lapse imaging of neuroblastoma cells to determine cell fate upon gene knockdown," *PLoS One*, vol. 7, no. 12, pp. e50988, 2012.
- [14] A. A. Taha and A. Hanbury, "Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool," *BMC Medical Imaging*, vol. 15(29), no. 1, 2015.
- [15] M. P. Dubuisson and A. K. Jain, "A modified Hausdorff distance for object matching," in *Pattern Recognition, 1994. Vol. 1 - Conference A: Computer Vision & Image Processing., Proc. 12th IAPR International Conference on*, Oct 1994, vol. 1, pp. 566–568 vol.1.