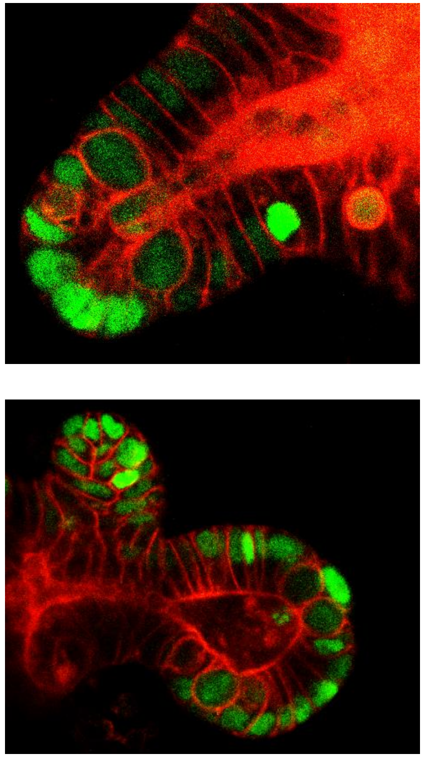
Mitosis Event Detection and Quantification

Student: Gerda Bortsova

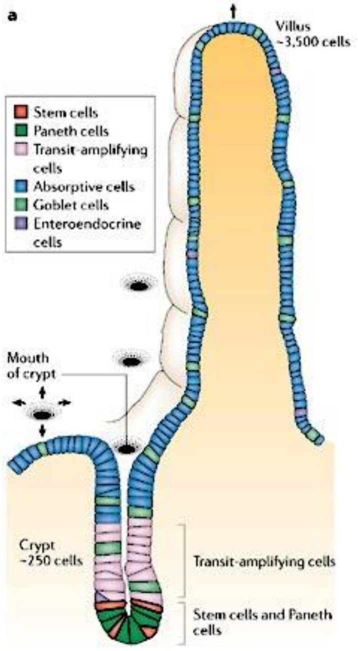
Supervisors: Dr. Lichao Wang, Dr. Carsten Marr

Introducing the problem I

- The data: 4D (3D + time) images of intestinal crypt (4 movies, ≈140 frames)
- Crypt is an invagination into mucosa of intestine [1]
- Where cells are actively proliferating to facilitate renewal of intestinal epithelium [3]
- Patterns of cell division **may be linked to diabetes**



The data: central slices.

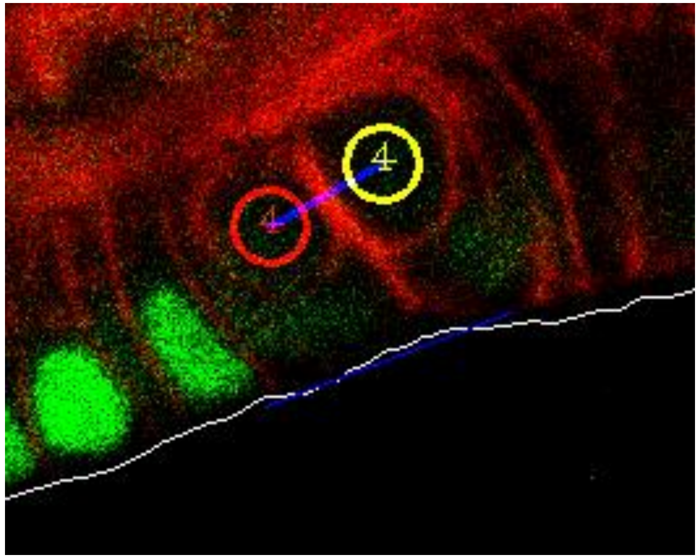


The distribution of epithelial cell types in the mammalian small intestine. [2]

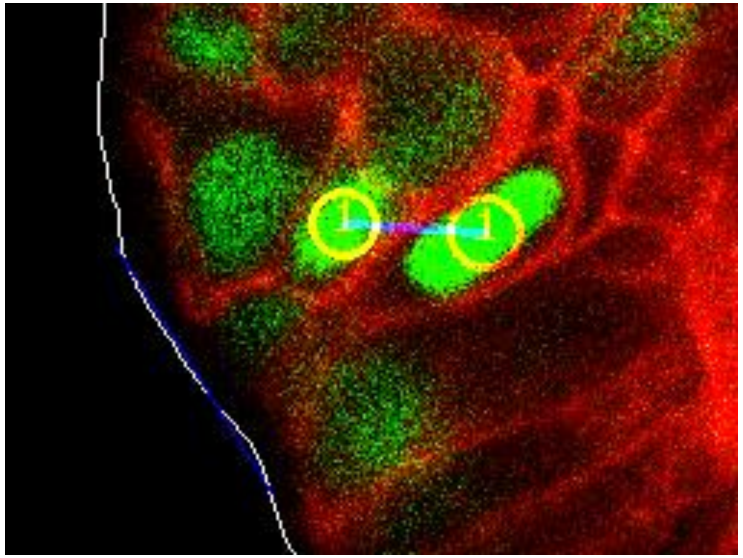
[1] Barker et al. 2008. The intestinal stem cell. *Genes & Dev.* **[2] Crosnier et at. 2006.** Organizing cell renewal in the intestine: stem cells, signals and combinatorial control. *Nature Reviews Genetics.* **[3] Yeung et al. 2011.** Regulation of self-renewal and differentiation by the intestinal stem cell niche. *Cellular and Molecular Life Sciences.*

Introducing the problem II

- **The task:** aid biologists to measure a statistics of angle between cells arisen from division and the crypt surface.



Almost parallel



Rather perpendicular

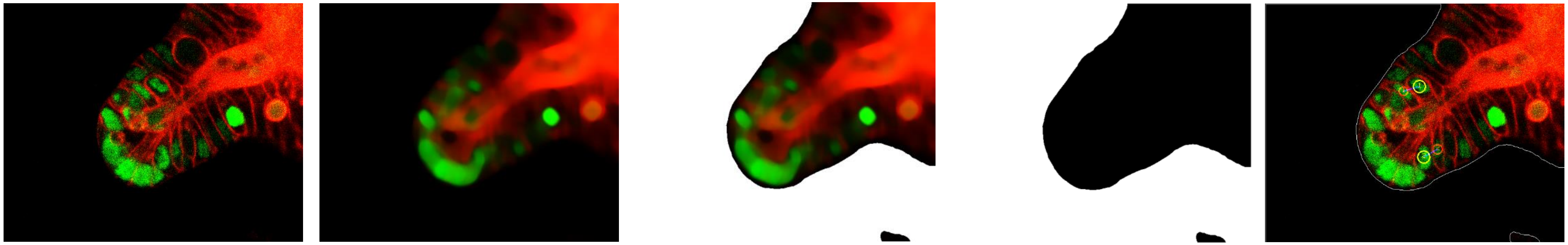
Introducing the problem III

The project parts:

- [ ] 1. Automatic extraction of crypt surface
- [ ] 2. Developing a method to measure angle between the crypt surface and daughter cells
- [ ] 3. Detection of mitosis events

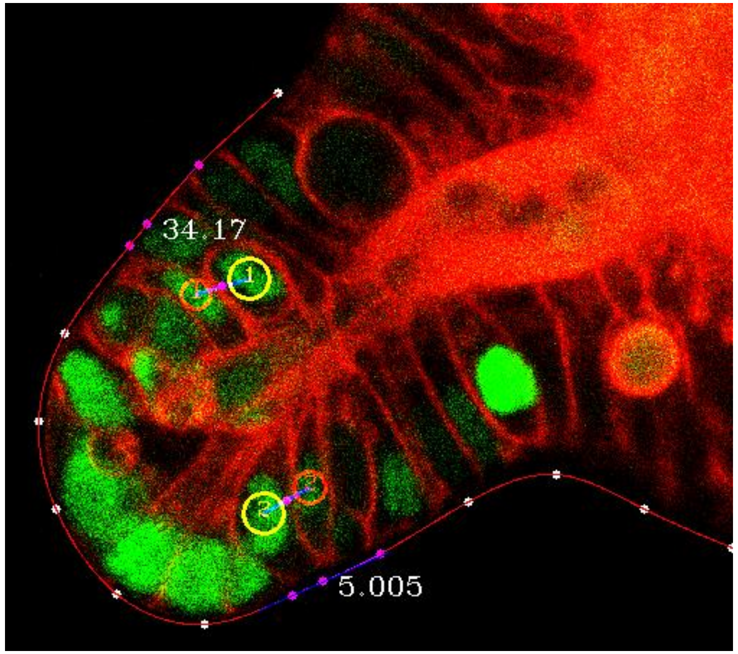
Extraction of crypt surface

- Statistical Shape Models (SSM) [4] was considered.
- Solved by Flood Fill algorithm [5] and subsequent extraction of the contour [6].

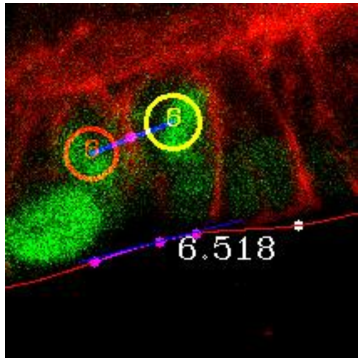


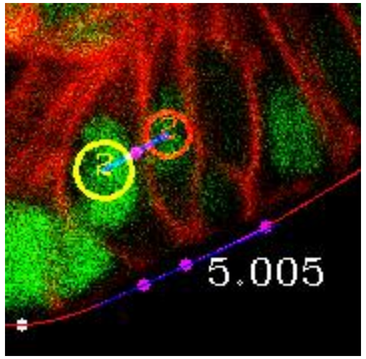
[4] Cootes et al. 2000. Statistical Models of Appearance for Computer Vision. **[5]** Flood Fill on Wikipedia: https://en.wikipedia.org/wiki/Flood_fill **[6] Suzuki. 1985.** Topological structural analysis of digitized binary images by border following. *Computer Vision, Graphics, and Image Processing.*

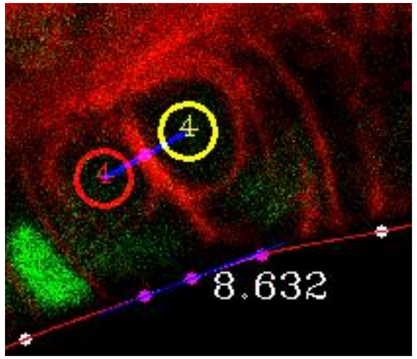
The method for measuring angle

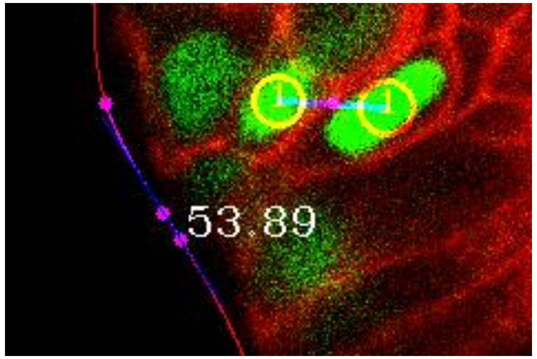


1. Sample points of equal distance from the crypt
2. Interpolate using cubic spline [7]
3. Calculate an angle between:
   - a line connecting daughter cells and
   - tangent line at the point of the crypt closest to a central point of a line segment connecting the daughters









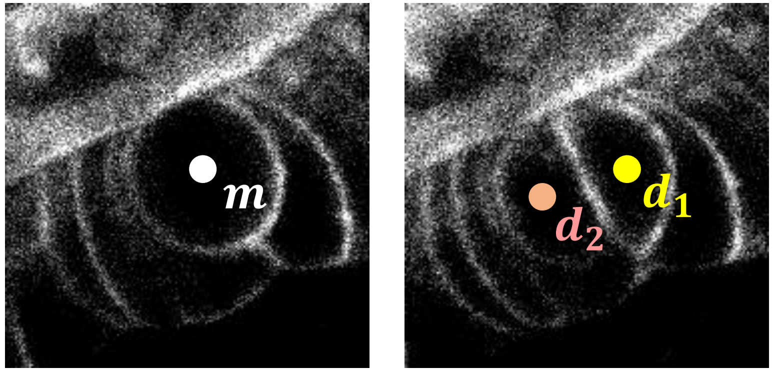
[7] Cubic Spline on Wolfram: http://mathworld.wolfram.com/CubicSpline.html

The project parts

- [x] 1. Automatic extraction of crypt surface
- [x] 2. Developing a method to measure angle between the crypt surface and daughter cells
- [ ] 3. Detection of mitosis events

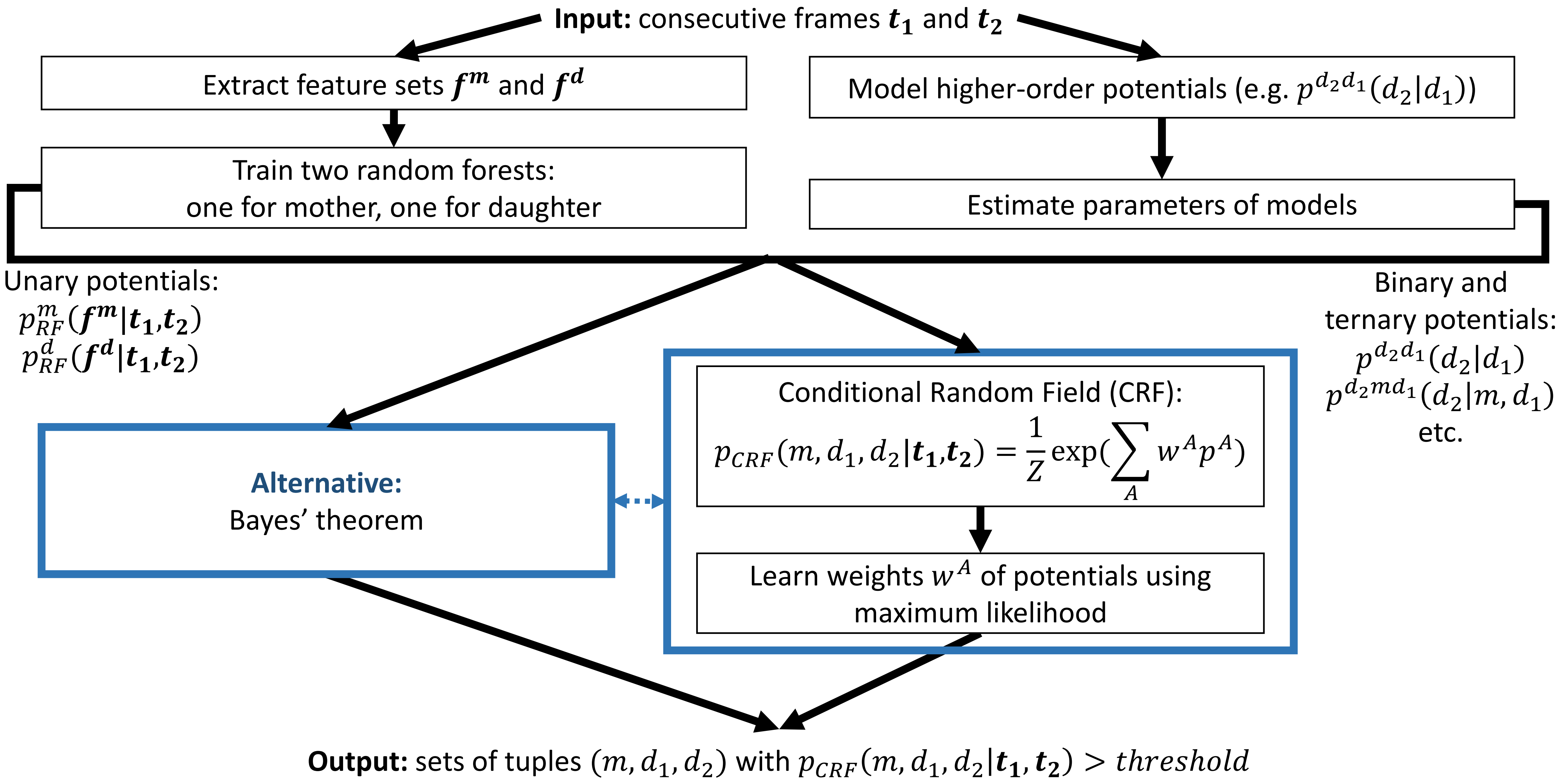
What is a mitosis event?

- A center of mother cell: m
- Centers of daughter cells arisen from the division of mother: d_1 and d_2

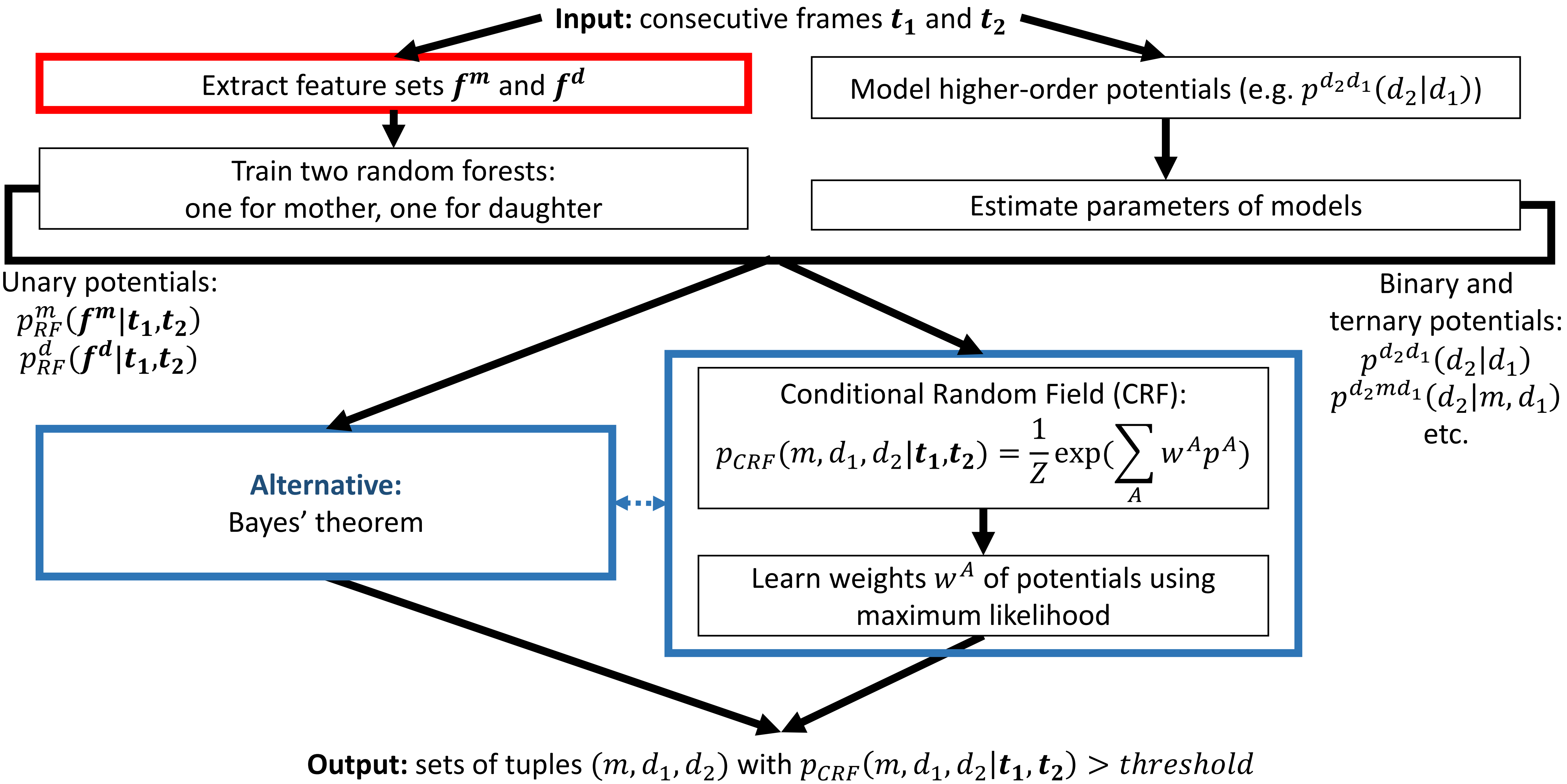


Time frame $\boldsymbol{t}$ Time frame $\boldsymbol{t+1}$

Detection of mitosis: the workflow

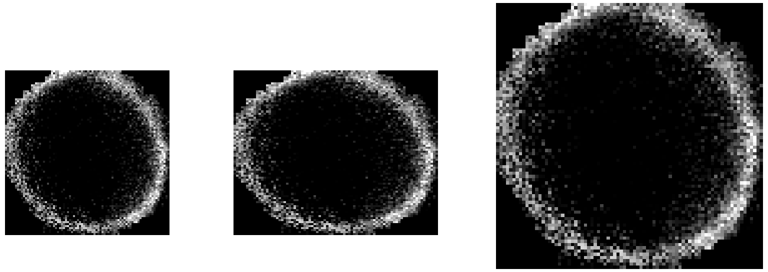


Detection of mitosis: the workflow

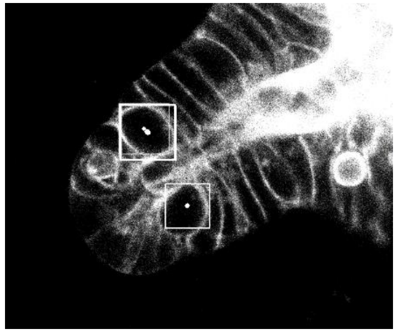


Features: template matching [8]

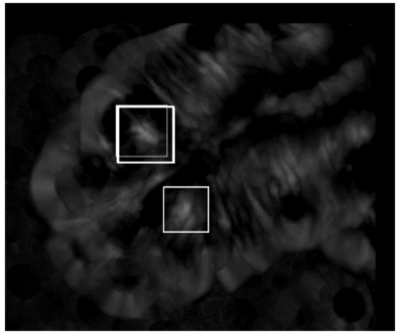
- The algorithm:
  - Choosing so-called template: image containing only the object
  - At each point of target image calculate a similarity measure (I took correlation coefficient) of a local patch and the template
- Problems:
  - Cells may have different sizes and not exactly circular shape -> search only for a single template, but for it’s linear transformations
  - Gives false positives (and negatives) -> use in combination with other features



Set of templates



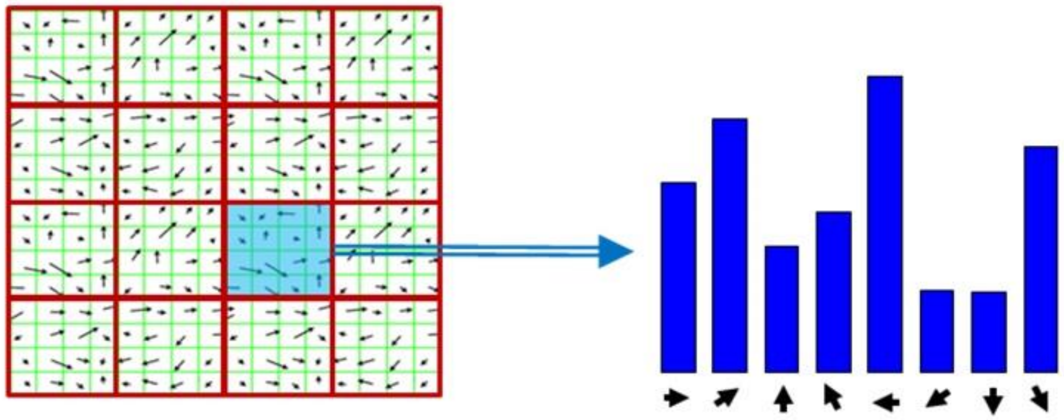
Mother cells detection



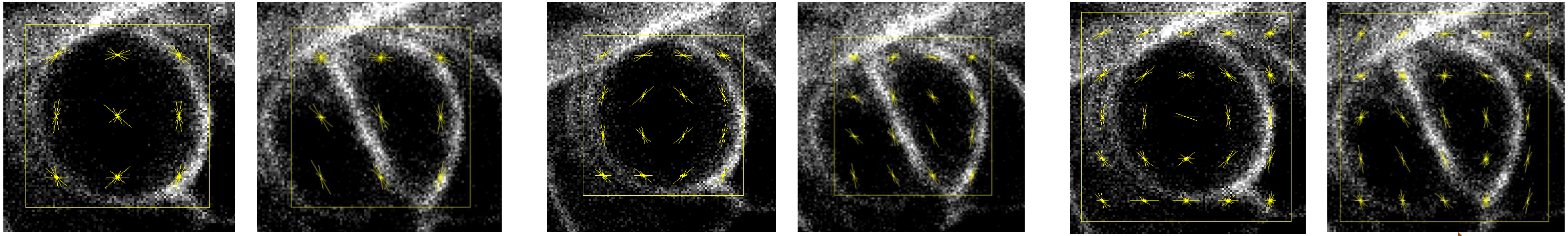
Correlation coefficient

[8] Brunelli. 2009. Template Matching Techniques in Computer Vision: Theory and Practice.

Features: histogram of oriented gradients (HOG) [9]



- The algorithm:
  - Compute gradient at each pixel of the patch
  - Each gradient votes for a bin in orientation-based histogram
- Problems:
  - Not rotation-invariant
  - Choose number of blocks:

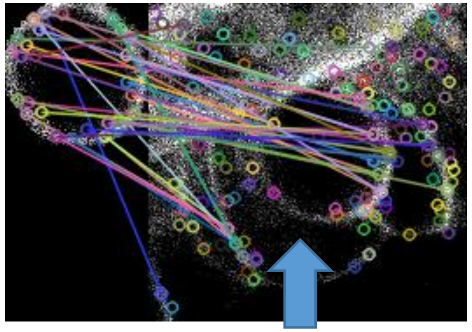


Number of parameters, Descriptiveness

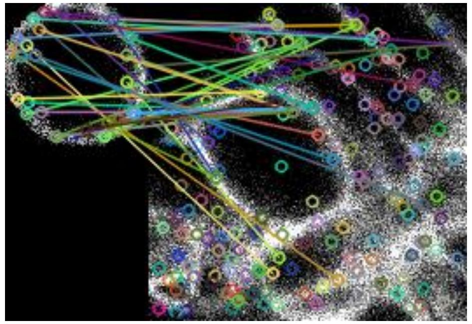
[9] Dalal et al. 2005. Histograms of oriented gradients for human detection. In Computer Vision and Pattern Recognition, 2005. *CVPR 2005*.

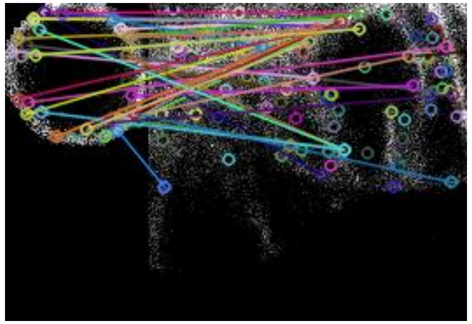
Features: SIFT (Scale-Invariant Feature Transform) [10]

- The algorithm:
  - Finds interesting points (keypoints) in both source and target images
  - Finds closest points on the target image to points on the source by comparing their HOGs
- Problems: is not suitable for this task, as cell is a homogeneous object



(Trying to find template on the image it was taken from.)





[10] Lowe. 2004. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision.*

Detection of mitosis: the workflow

Input: consecutive frames $\boldsymbol{t_1}$ and $\boldsymbol{t_2}$

Extract feature sets $\boldsymbol{f^m}$ and $\boldsymbol{f^d}$

Train two random forests:
one for mother, one for daughter

Model higher-order potentials (e.g. $p^{d_2 d_1}(d_2|d_1)$)

Estimate parameters of models

Unary potentials:
$p_{RF}^{m}(\boldsymbol{f^m}|\boldsymbol{t_1},\boldsymbol{t_2})$
$p_{RF}^{d}(\boldsymbol{f^d}|\boldsymbol{t_1},\boldsymbol{t_2})$

Binary and
ternary potentials:
$p^{d_2 d_1}(d_2|d_1)$
$p^{d_2 m d_1}(d_2|m,d_1)$
etc.

Alternative:
Bayes' theorem

Conditional Random Field (CRF):

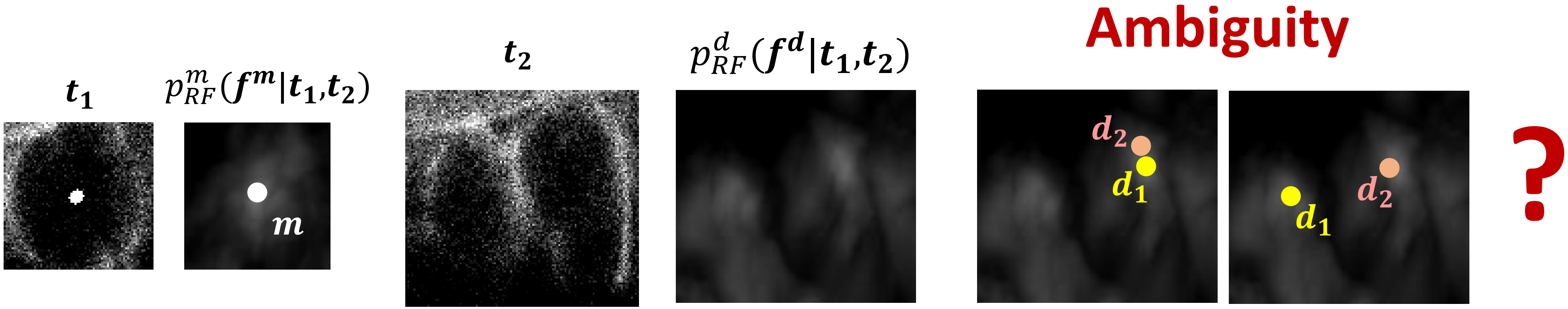
$$p_{CRF}(m, d_1, d_2|\boldsymbol{t_1},\boldsymbol{t_2}) = \frac{1}{Z}\exp(\sum_A w^A p^A)$$

Learn weights w^A of potentials using
maximum likelihood

Output: sets of tuples (m, d_1, d_2) with $p_{CRF}(m, d_1, d_2|\boldsymbol{t_1},\boldsymbol{t_2}) > threshold$

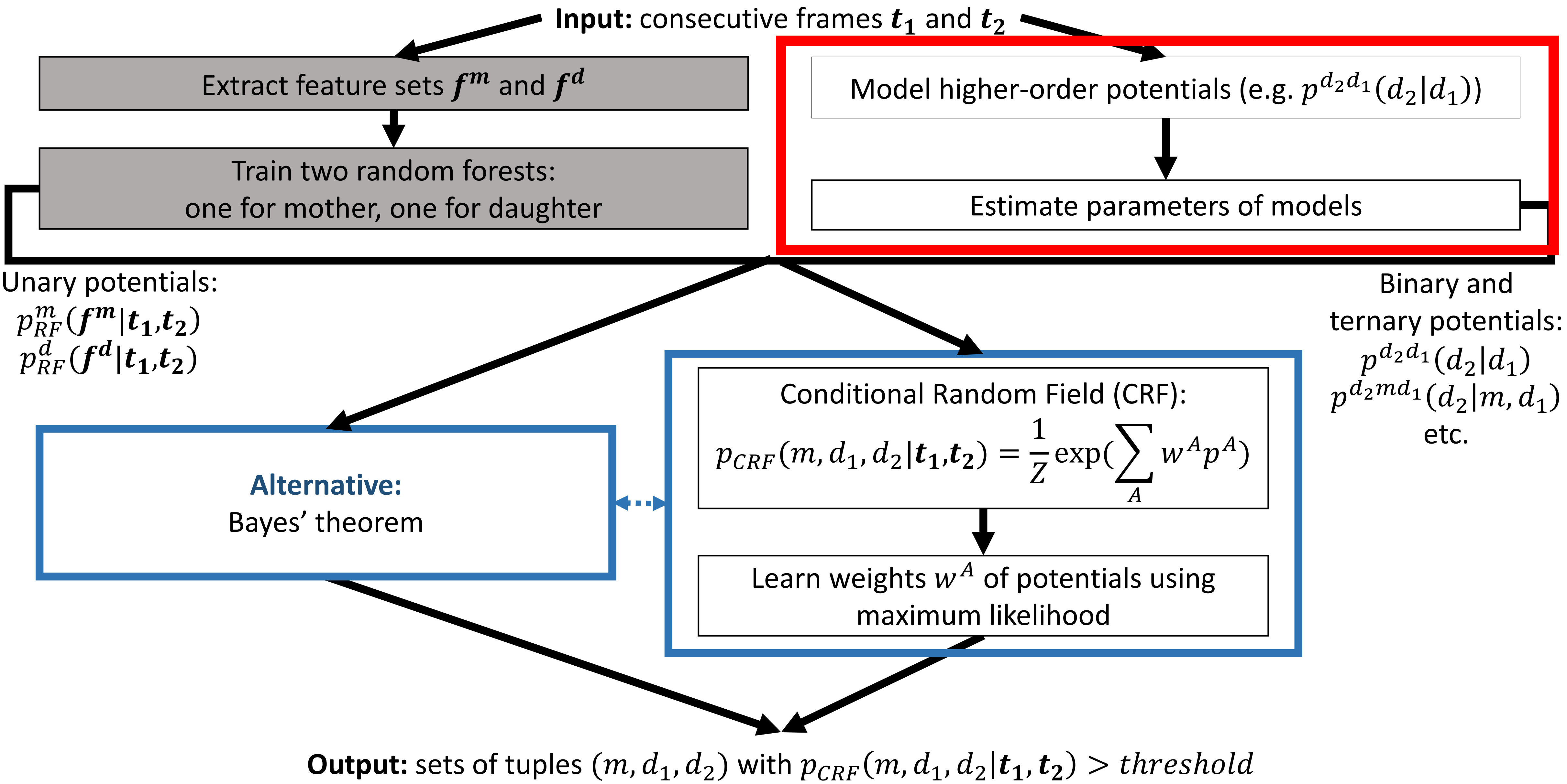
Training classifier to learn unary potentials

- Extract discriminative features: template matching, HOG, color-based features
- Train two random forest [11] classifiers: for mother-background and for daughter-background classification



[11] Bishop 2006. Pattern Recognition and Machine Learning.

Detection of mitosis: the workflow

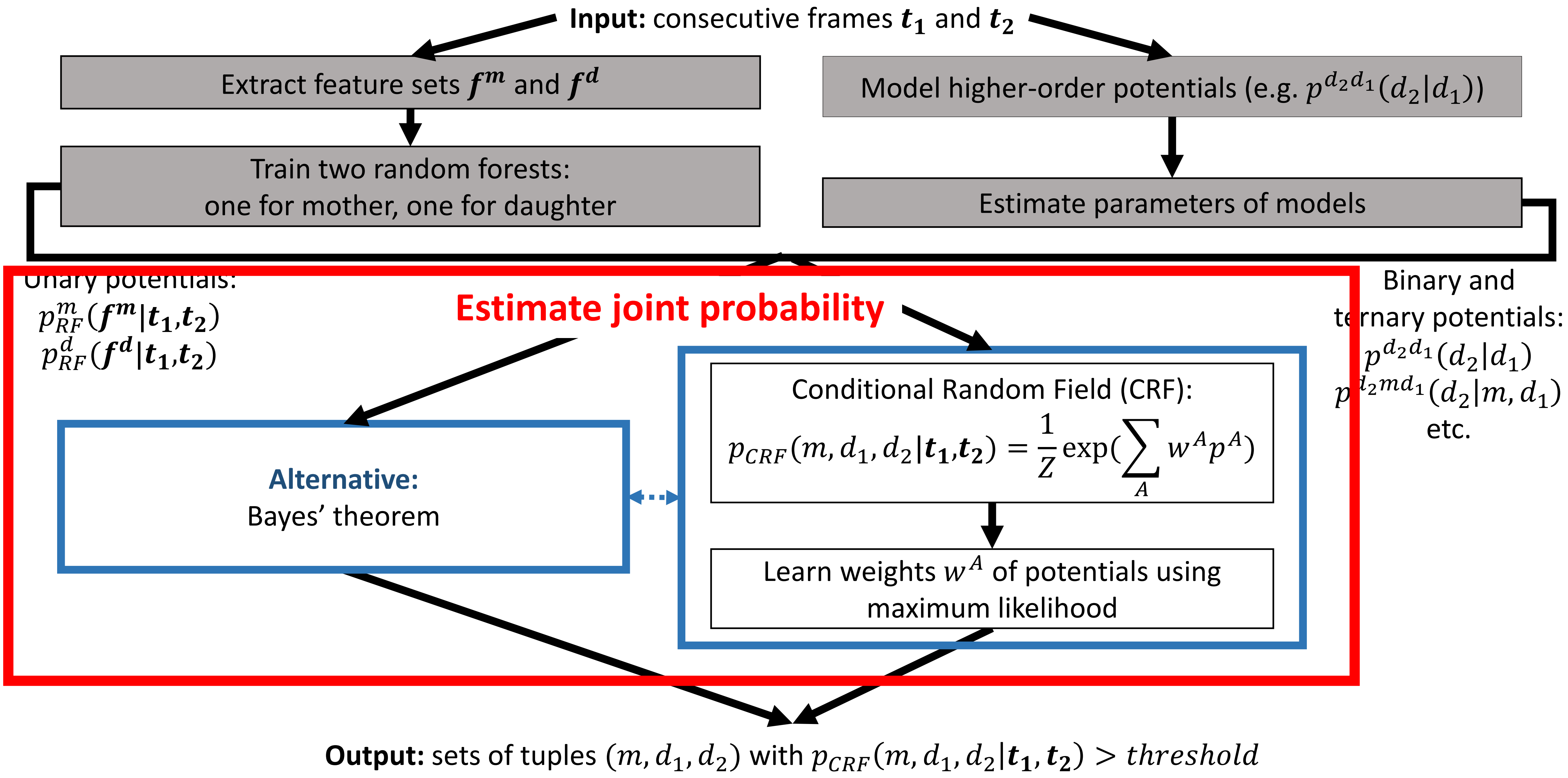


Higher-order potentials modeling

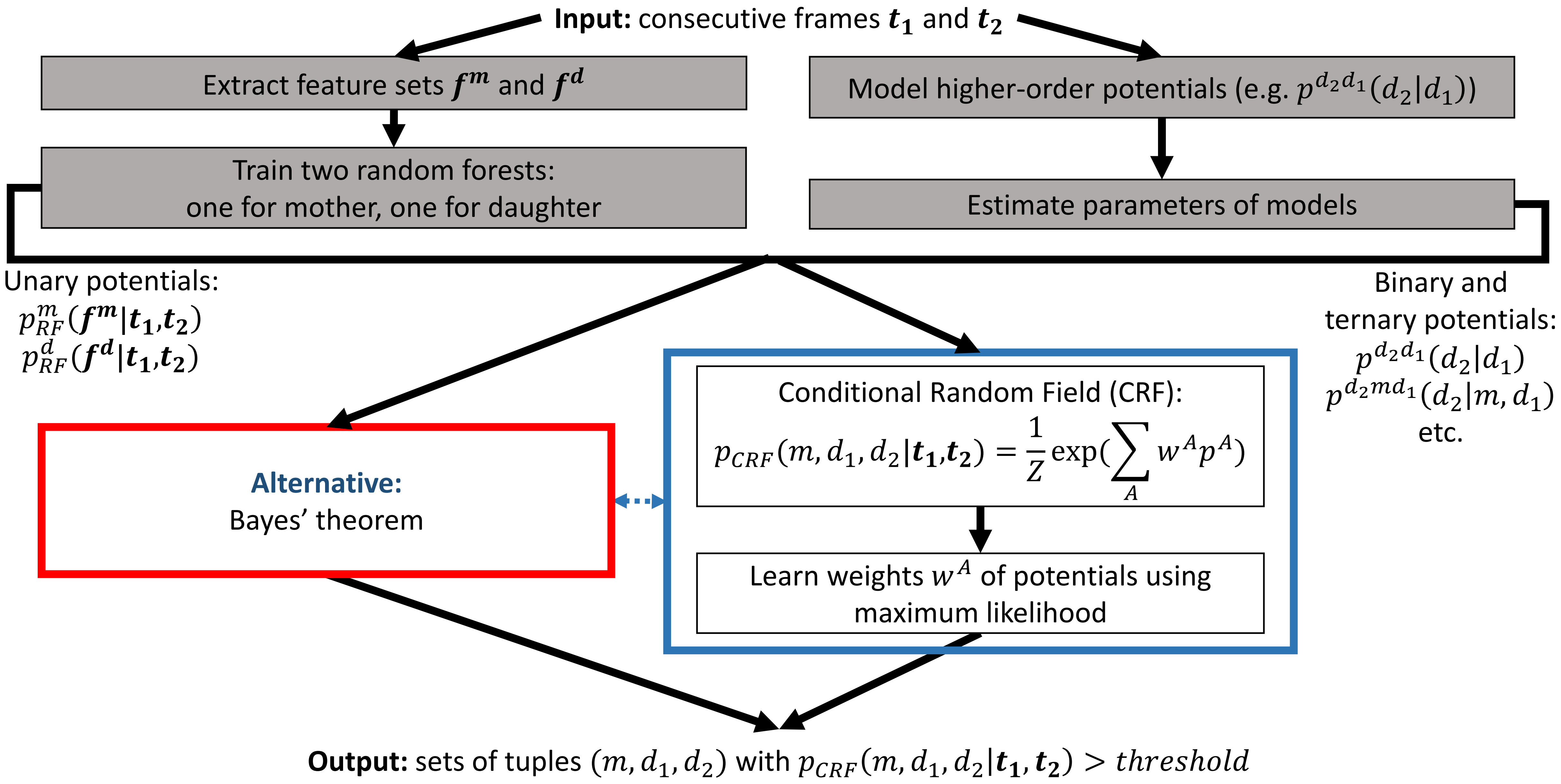
- For binary potentials: transition is a Gaussian
  - E.g. $p(d_2|d_1) = N(|d_2 - d_1| \mid \mu, \sigma)$
- For ternary potentials: change in transition
  - E.g. $p(d_2|m, d_1) = N\big(|m - d_1| - |m - d_2| \big| \mu, \sigma\big)$

- Parameters μ and σ estimated from training set.

Detection of mitosis: the workflow



Detection of mitosis: the workflow



Estimating joint probability: Bayes

- Employing probability :

$$p(m, d_1, d_2) = p(m)p(d_1)p(d_2|d_1)$$

$$p_{RF}^{m}(\boldsymbol{f^m}(m)|\boldsymbol{t_1},\boldsymbol{t_2}) \qquad p_{RF}^{d}\left(\boldsymbol{f^d}(d)|\boldsymbol{t_1},\boldsymbol{t_2}\right) \qquad N(|d_2 - d_1| \mid \mu, \sigma)$$

- Drawbacks:
  - probability estimates could be inaccurate
  - restricted in number of potentials to incorporate

Detection of mitosis: the workflow

Input: consecutive frames $\boldsymbol{t_1}$ and $\boldsymbol{t_2}$

Extract feature sets $\boldsymbol{f^m}$ and $\boldsymbol{f^d}$

Train two random forests:
one for mother, one for daughter

Model higher-order potentials (e.g. $p^{d_2 d_1}(d_2|d_1)$)

Estimate parameters of models

Unary potentials:
$p_{RF}^{m}(\boldsymbol{f^m}|\boldsymbol{t_1},\boldsymbol{t_2})$
$p_{RF}^{d}(\boldsymbol{f^d}|\boldsymbol{t_1},\boldsymbol{t_2})$

Binary and
ternary potentials:
$p^{d_2 d_1}(d_2|d_1)$
$p^{d_2 m d_1}(d_2|m, d_1)$
etc.

Conditional Random Field (CRF):

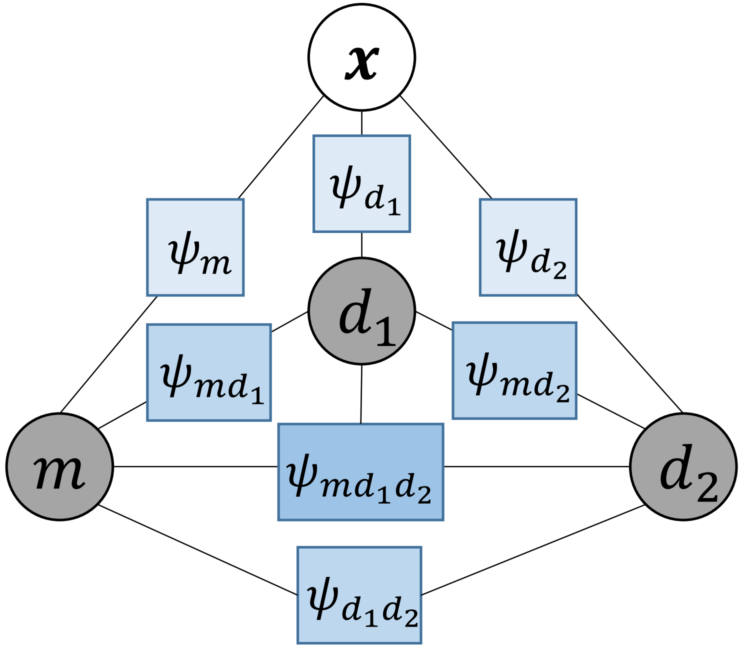
$$p_{CRF}(m, d_1, d_2|\boldsymbol{t_1},\boldsymbol{t_2}) = \frac{1}{Z}\exp(\sum_A w^A p^A)$$

Learn weights w^A of potentials using
maximum likelihood

Alternative:
Bayes' theorem

Output: sets of tuples (m, d_1, d_2) with $p_{CRF}(m, d_1, d_2|\boldsymbol{t_1}, \boldsymbol{t_2}) > threshold$

Conditional Random Field (CRF) formulation [12, 13]



- $U = \{m, d_1, d_2\}$
- $B = \{(m, d_1), (m, d_2), (d_1, d_2)\}$
- $T = \{(m, d_1, d_2)\}$

$$p(m, d_1, d_2|\boldsymbol{x}) = \frac{1}{Z(x)} \prod_{i \in U} \psi_i(i, \boldsymbol{x}) \prod_{(i,j) \in B} \psi_{ij}(i, j) \prod_{(i,j,k) \in T} \psi_{ijk}(i, j, k)$$

[12] Lafferty et al. 2001. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. *International Conference on Machine Learning.* **[13] Sutton et al. 2010.** An Introduction to Conditional Random Fields.

CRF formulation: factors

- Unary factors:

$$\psi_m(m, x) = \exp\{w_m p(m)\}$$

- Binary factors:

$$\psi_{d_1 d_2}(d_1, d_2, x) = \exp\{w_{d_1 d_2} p(d_1 | d_2)\}$$

- Ternary factors:

$$\psi_{m d_1 d_2}(m, d_1, d_2, x) = \exp\{w_{m d_1 d_2} p(m | d_1, d_2)\}$$

CRF formulation: learn weights

- In some works weights are chosen empirically. [14]
- Disadvantages:
  - Requires time to tune
  - May be inaccurate
- Better option: learn weights by maximizing likelihood [15]:

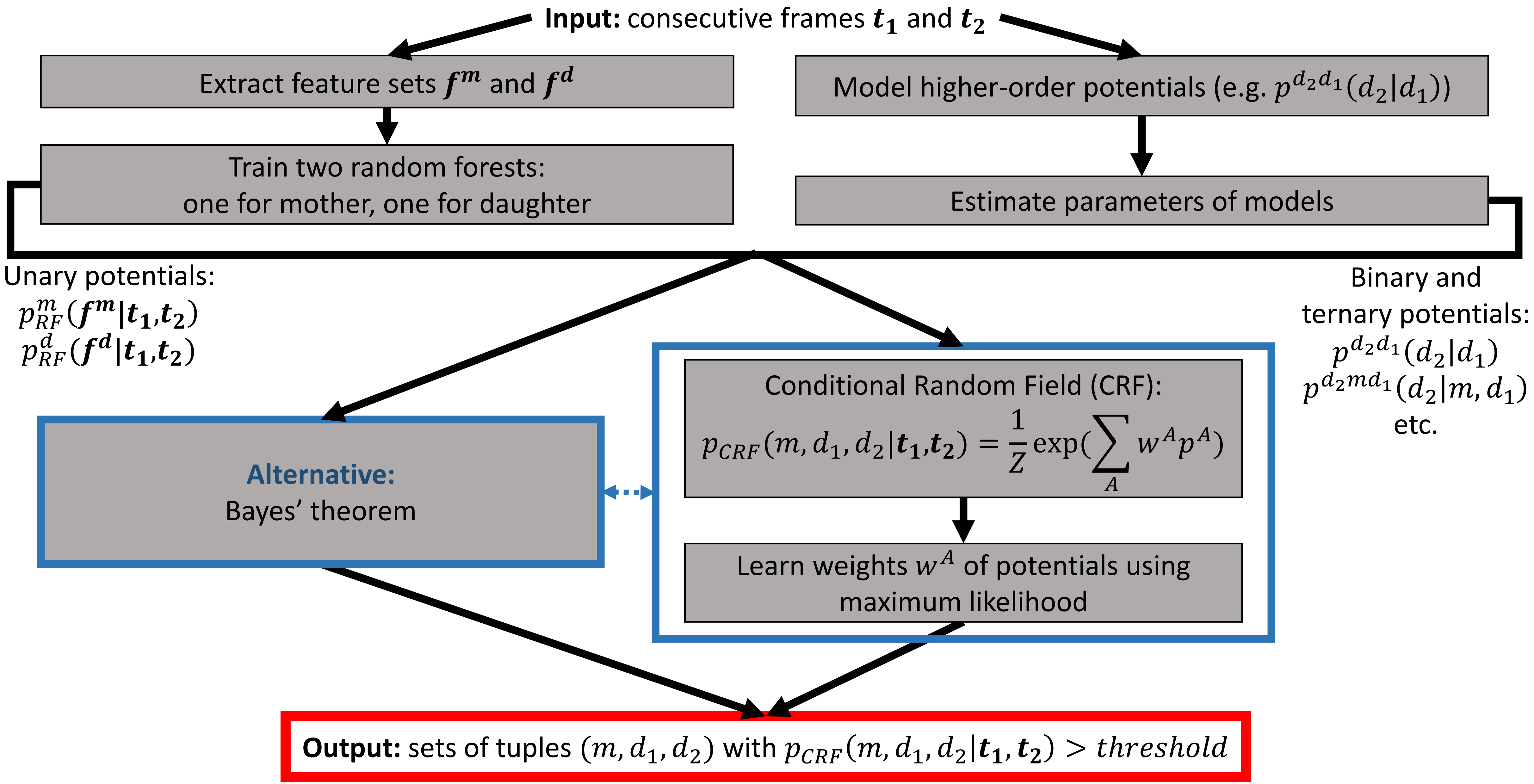
$$\widehat{w} = \arg max_w \prod_{\{n=1\}}^{N} p(m^n, d_1^n, d_1^n | \boldsymbol{x^n}) = \prod_{\{n=1\}}^{N} \frac{1}{Z} \exp(\sum_{A} w^A p_n^A)$$

- m^n, d_1^n, d_1^n - labels of nth sample $\boldsymbol{x^n}$
- Positive samples: annotated cell divisions; negative samples: random triples

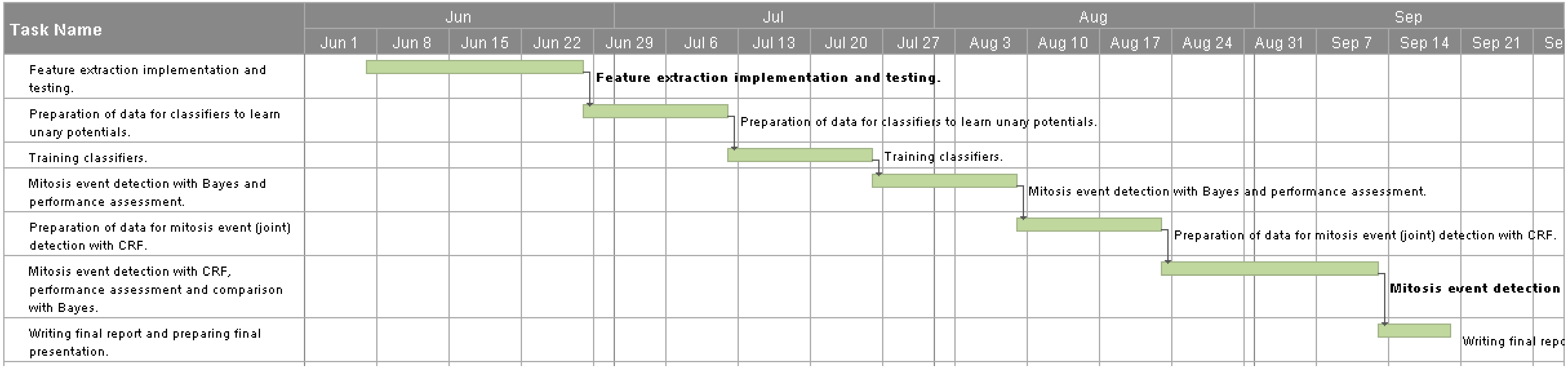
[14] Wang et al. 2013. Landmark detection and coupled patch registration for cardiac motion tracking. *SPIE.*

[15] Wang et al. 2015. Anatomic-Landmark Detection Using Graphical Context Modelling.

Detection of mitosis: the workflow



The schedule



References

[1] Barker, N., van de Wetering, M., and Clevers, H. 2008. The intestinal stem cell. Genes & Dev. 22:1856-1864.

[2] Crosnier, C., Stamataki, D., Lewis, J. 2006. Organizing cell renewal in the intestine: stem cells, signals and combinatorial control. Nature Reviews Genetics. 7(5):349-359.

[3] Yeung, T.M., Chia, L.A., Kosinski, C.M., Kuo, C.J. 2011. Regulation of self-renewal and differentiation by the intestinal stem cell niche. Cellular and Molecular Life Sciences. Vol. 68, no. 15, pp. 2513–2523.

[4] Cootes, T.F., Taylor, C. J. 2000. Statistical Models of Appearance for Computer Vision.

[5] Flood fill Wikipedia: https://en.wikipedia.org/wiki/Flood_fill

[6] Suzuki, S. (1985). Topological structural analysis of digitized binary images by border following. Computer Vision, Graphics, and Image Processing, 30(1), 32-46.

[7] Cubic Spline on Wolfram: http://mathworld.wolfram.com/CubicSpline.html

[8] Template Matching Techniques in Computer Vision: Theory and Practice. Roberto Brunelli. March 2009. ISBN: 978-0-470-51706-2

[9] Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on (Vol. 1, pp. 886-893). IEEE.

[10] Distinctive image features from scale-invariant keypoints. David G. Lowe, International Journal of Computer Vision, 60, 2 (2004), pp. 91-110.

[11] Christopher M. Bishop, Pattern Recognition and Machine Learning. Springer, Berlin, New York, 2006.

[12] Lafferty, J., McCallum, A., Pereira, F. 2001. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. Proceedings of 18th International Conference on Machine Learning.

[13] Sutton, C., McCallum, A. 2010. An Introduction to Conditional Random Fields. v1. arXiv:1011.4088.

[14] H. Wang, W. Shi, X. Zhuang, X. Wu, K. Tung, S. Ourselin, P. Edwards, and D. Rueckert, "Landmark detection and coupled patch registration for cardiac motion tracking," in SPIE, 2013, vol. 8669.

[15] Wang, L., Belagiannis, V., Marr, C., Theis, F., Yang, G., Navab, N. 2015. Anatomic-Landmark Detection Using Graphical Context Modelling.