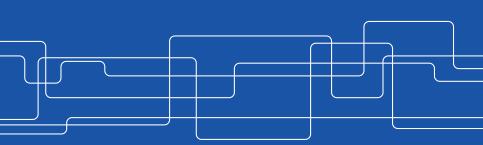


Cell detection by functional inverse diffusion and non-negative group sparsity

Pol del Aguila Pla, Ph.D. Candidate https://people.kth.se/~poldap Division of Information Science and Engineering School of Electrical Engineering and Computer Science

January 14, 2019 at 🌀 CosmoStat



KTH Royal Institute of Technology

Acknowledgements



J. Jaldén [1]–[5] V. Saxena [5] G. Bengtsson J. Larsson J. Sörell E. Ågeby



Royal Swedish Academy of Sciences



Swedish Research Council

Knut and Alice Wallenberg foundation

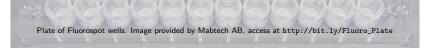


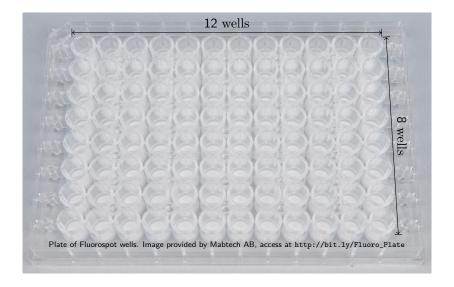
KTH Opportunities and EECS school

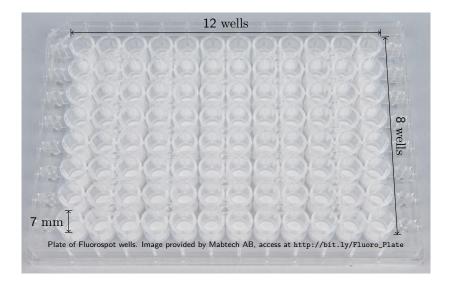


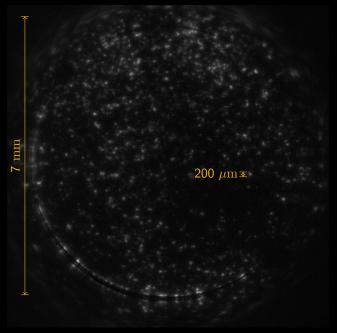
MabTech AB

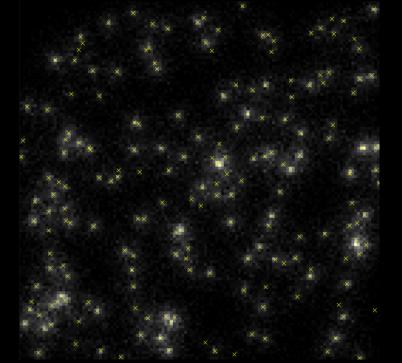
- P. del Aguila Pla and J. Jaldén, "Cell detection by functional inverse diffusion and non-negative group sparsity—Part I: Modeling and Inverse problems," *IEEE Transactions on Signal Processing*, vol. 66, no. 20, pp. 5407–5421, Oct. 2018
- P. del Aguila Pla and J. Jaldén, "Cell detection by functional inverse diffusion and non-negative group sparsity—Part II: Proximal optimization and Performance evaluation," *IEEE Transactions on Signal Processing*, vol. 66, no. 20, pp. 5422–5437, Oct. 2018
- P. del Aguila Pla and J. Jaldén, "Cell detection on image-based immunoassays," in 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI), Apr. 2018, pp. 431–435
- P. del Aguila Pla and J. Jaldén, "Convolutional group-sparse coding and source localization," in 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Apr. 2018, pp. 2776–2780
- P. del Aguila Pla, V. Saxena, and J. Jaldén, "SpotNet Learned iterations for cell detection in image-based immunoassays,", Accepted in 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI), Apr. 2019

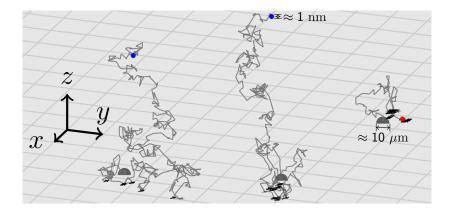












Relevant quantities for the assay are

A density of bound particles d(x, y, t) ≥ 0, where the image will be d_{obs}(x, y) = d(x, y, T), which evolves coupled to

Relevant quantities for the assay are

- A density of bound particles d(x, y, t) ≥ 0, where the image will be d_{obs}(x, y) = d(x, y, T), which evolves coupled to
- ► the 3D density of free particles c(x, y, z, t) ≥ 0 on z ≥ 0, and to

Relevant quantities for the assay are

- A density of bound particles d(x, y, t) ≥ 0, where the image will be d_{obs}(x, y) = d(x, y, T), which evolves coupled to
- ► the 3D density of free particles c(x, y, z, t) ≥ 0 on z ≥ 0, and to
- ► the source density rate of new particles s(x, y, t) ≥ 0, that is spatially sparse and reveals the cell locations and secretion over time.

Relevant quantities for the assay are

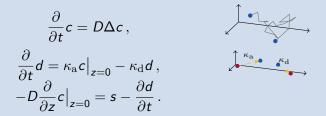
- A density of bound particles d(x, y, t) ≥ 0, where the image will be d_{obs}(x, y) = d(x, y, T), which evolves coupled to
- ► the 3D density of free particles c(x, y, z, t) ≥ 0 on z ≥ 0, and to
- ► the source density rate of new particles s(x, y, t) ≥ 0, that is spatially sparse and reveals the cell locations and secretion over time.

$$\begin{split} \frac{\partial}{\partial t} c &= D\Delta c \,, \\ \frac{\partial}{\partial t} d &= \kappa_{\rm a} c \big|_{z=0} - \kappa_{\rm d} d \\ - D \frac{\partial}{\partial z} c \big|_{z=0} &= s - \frac{\partial d}{\partial t} \end{split}$$

This physical model was presented before, also for ELISPOT and Fluorospot.

Relevant quantities for the assay are

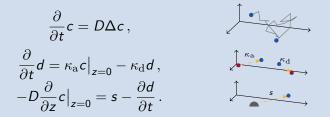
- A density of bound particles d(x, y, t) ≥ 0, where the image will be d_{obs}(x, y) = d(x, y, T), which evolves coupled to
- ▶ the 3D density of free particles $c(x, y, z, t) \ge 0$ on $z \ge 0$, and to
- ► the source density rate of new particles s(x, y, t) ≥ 0, that is spatially sparse and reveals the cell locations and secretion over time.



This physical model was presented before, also for ELISPOT and Fluorospot.

Relevant quantities for the assay are

- A density of bound particles d(x, y, t) ≥ 0, where the image will be d_{obs}(x, y) = d(x, y, T), which evolves coupled to
- ▶ the 3D density of free particles $c(x, y, z, t) \ge 0$ on $z \ge 0$, and to
- ► the source density rate of new particles s(x, y, t) ≥ 0, that is spatially sparse and reveals the cell locations and secretion over time.



This physical model was presented before, also for ELISPOT and Fluorospot.

We consider the image observation $d_{\rm obs} \in \mathcal{D}_+$, with $\mathcal{D} = L^2(\mathbb{R}^2)$ and prove that

$$d_{\rm obs}(x,y) = \int_0^{\sigma_{\rm max}} \left(g_{\sigma}(\bar{x},\bar{y}) * a(\bar{x},\bar{y},\sigma) \right)(x,y) \, \mathrm{d}\sigma \, ,$$

with $a \in A_+$ and $A \subset L^2(\mathbb{R}^2 \times \mathbb{R}_+)$ a space of functions with bounded spatial support, $\sigma_{\max} = \sqrt{2DT}$, and

We consider the image observation $d_{\rm obs} \in \mathcal{D}_+$, with $\mathcal{D} = L^2(\mathbb{R}^2)$ and prove that

$$d_{
m obs}(x,y) = \int_0^{\sigma_{
m max}} G_\sigma a_\sigma \,\mathrm{d}\sigma\,,$$

with $a \in A_+$ and $A \subset L^2(\mathbb{R}^2 \times \mathbb{R}_+)$ a space of functions with bounded spatial support, $\sigma_{\max} = \sqrt{2DT}$, and

We consider the image observation $d_{\rm obs}\in\mathcal{D}_+$, with $\mathcal{D}={
m L}^2\left(\mathbb{R}^2
ight)$ and prove that

 $d_{obs}(x, y) = Aa$, we call A the *diffusion* operator,

with $a \in A_+$ and $A \subset L^2(\mathbb{R}^2 \times \mathbb{R}_+)$ a space of functions with bounded spatial support, $\sigma_{\max} = \sqrt{2DT}$, and

We consider the image observation $d_{\rm obs}\in\mathcal{D}_+$, with $\mathcal{D}={
m L}^2\left(\mathbb{R}^2
ight)$ and prove that

 $d_{obs}(x, y) = Aa$, we call A the *diffusion* operator,

with $a \in A_+$ and $A \subset L^2(\mathbb{R}^2 \times \mathbb{R}_+)$ a space of functions with bounded spatial support, $\sigma_{\max} = \sqrt{2DT}$, and

 a(x, y, σ) is an equivalent of s(x, y, t) where the effect of adsorption and desorption have been summarized.

$$a(x, y, \sigma) = \frac{\sigma}{D} \int_{\frac{\sigma^2}{2D}}^{T} s(x, y, T - \eta) \varphi\left(\frac{\sigma^2}{2D}, \eta\right) \mathrm{d}\eta.$$

• $a(x, y, \sigma)$ preserves all the spatial information in s(x, y, t).

The modeling result: The image $d_{\mathrm{obs}} \in \mathcal{D}_+$ can be expressed as $\ell^{\sigma_{\mathsf{max}}}$

$$d_{\rm obs} = \int_0 \quad G_\sigma a_\sigma \mathrm{d}\sigma \,.$$

How?



Independence of Brownian motion in x, y and z.

The modeling result: The image $d_{\mathrm{obs}} \in \mathcal{D}_+$ can be expressed as $\ell^{\sigma_{\max}}$

$$d_{\rm obs} = \int_0^{-\infty} G_\sigma a_\sigma {\rm d}\sigma \,.$$

- Independence of Brownian motion in x, y and z.
- Adsorption (κ_a) and desorption (κ_d) only regulated by z-movement.

The modeling result: The image $d_{\mathrm{obs}} \in \mathcal{D}_+$ can be expressed as $\ell^{\sigma_{\max}}$

$$d_{\rm obs} = \int_0^{-\infty} G_\sigma a_\sigma \mathrm{d}\sigma \,.$$

- Independence of Brownian motion in x, y and z.
- Adsorption (κ_a) and desorption (κ_d) only regulated by z-movement.
- ► x- and y-movements only depend on τ , total time in Brownian motion. In particular, according to Green function for 2D diffusion, $g_{\sqrt{2D\tau}}(x, y)$.

The modeling result: The image $d_{\mathrm{obs}} \in \mathcal{D}_+$ can be expressed as $\ell^{\sigma_{\max}}$

$$d_{\rm obs} = \int_0^{-\infty} G_\sigma a_\sigma \mathrm{d}\sigma \,.$$

- Independence of Brownian motion in x, y and z.
- Adsorption (κ_a) and desorption (κ_d) only regulated by z-movement.
- ► x- and y-movements only depend on τ , total time in Brownian motion. In particular, according to Green function for 2D diffusion, $g_{\sqrt{2D\tau}}(x, y)$.
- φ(τ, t) summarizes the effect of adsorption and desorption onto the time in free motion τ for each time of final adsorption t.

The modeling result: The image $d_{\mathrm{obs}} \in \mathcal{D}_+$ can be expressed as $\ell^{\sigma_{\max}}$

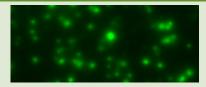
$$d_{\rm obs} = \int_0^{-\infty} G_\sigma a_\sigma {\rm d}\sigma \,.$$

- Independence of Brownian motion in x, y and z.
- Adsorption (κ_a) and desorption (κ_d) only regulated by z-movement.
- ► x- and y-movements only depend on τ , total time in Brownian motion. In particular, according to Green function for 2D diffusion, $g_{\sqrt{2D\tau}}(x, y)$.
- φ(τ, t) summarizes the effect of adsorption and desorption onto the time in free motion τ for each time of final adsorption t.
- Change variables to those significative to x- and y-movement, $\sigma = \sqrt{2D\tau}$.

The modeling result: The image $d_{\mathrm{obs}} \in \mathcal{D}_+$ can be expressed as

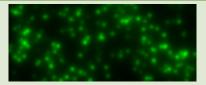
$$d_{\rm obs} = \int_0^{T_{\rm omax}} G_\sigma a_\sigma {\rm d}\sigma \,.$$

Consequences



Real observation (section)

Synthetic data

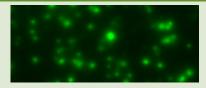


Simulated observation (section)

The modeling result: The image $d_{\mathrm{obs}} \in \mathcal{D}_+$ can be expressed as

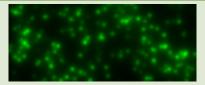
$$d_{\rm obs} = \int_0^{\sigma_{\rm max}} G_\sigma a_\sigma {\rm d}\sigma \,.$$

Consequences



Real observation (section)

- Synthetic data
- An inverse problem



Simulated observation (section)

We have $d_{obs} \in D_+$ and want to recover $a \in A_+$. We propose the (non-smooth, constrained) convex problem

$$\min_{a \in \mathcal{A}} \left[\|Aa - d_{obs}\|_{\mathcal{D}}^{2} + \underbrace{\delta_{\mathcal{A}_{+}}(a)}_{\text{non-negative}} + \lambda \underbrace{\int_{\mathbb{R}^{2}} \left(\int_{0}^{\sigma_{max}} a^{2}(x, y, \sigma) \, \mathrm{d}\sigma \right)^{\frac{1}{2}} \mathrm{d}x \mathrm{d}y}_{\text{group-sparsity}} \right],$$

We have $d_{obs} \in D_+$ and want to recover $a \in A_+$. We propose the (non-smooth, constrained) convex problem

$$\min_{\boldsymbol{a}\in\mathcal{A}} \left[\left\| \boldsymbol{A}\boldsymbol{a} - \boldsymbol{d}_{\mathrm{obs}} \right\|_{\mathcal{D}}^{2} + \delta_{\mathcal{A}_{+}}(\boldsymbol{a}) + \lambda \left\| \left\| \boldsymbol{a}_{\mathsf{r}} \right\|_{\mathrm{L}^{2}(\mathbb{R}_{+})} \right\|_{\mathrm{L}^{1}(\mathbb{R}^{2})} \right]$$

with $\mathbf{r} = (x, y)$.

We have $d_{obs} \in D_+$ and want to recover $a \in A_+$. We propose the (non-smooth, constrained) convex problem

$$\min_{\boldsymbol{a}\in\mathcal{A}} \left[\left\| \boldsymbol{A}\boldsymbol{a} - \boldsymbol{d}_{\mathrm{obs}} \right\|_{\mathcal{D}}^{2} + \delta_{\mathcal{A}_{+}}(\boldsymbol{a}) + \lambda \left\| \left\| \boldsymbol{a}_{\mathsf{r}} \right\|_{\mathrm{L}^{2}(\mathbb{R}_{+})} \right\|_{\mathrm{L}^{1}(\mathbb{R}^{2})} \right]$$

with $\mathbf{r} = (x, y)$.

Proximal Optimization

How do we solve this optimization problem? Can it be solved?

We have $d_{obs} \in D_+$ and want to recover $a \in A_+$. We propose the (non-smooth, constrained) convex problem

$$\min_{\mathbf{a}\in\mathcal{A}} \left[\left\| A\mathbf{a} - d_{\mathrm{obs}} \right\|_{\mathcal{D}}^{2} + \delta_{\mathcal{A}_{+}}(\mathbf{a}) + \lambda \left\| \left\| \mathbf{a}_{\mathsf{r}} \right\|_{\mathrm{L}^{2}(\mathbb{R}_{+})} \right\|_{\mathrm{L}^{1}(\mathbb{R}^{2})} \right]$$

with $\mathbf{r} = (x, y)$.

- How do we solve this optimization problem? Can it be solved?
- Three terms, two non-smooth (with known prox), one smooth (with non-trivial but manageable gradient). Convex problem, but existance and unicity not given (function spaces).

We have $d_{obs} \in D_+$ and want to recover $a \in A_+$. We propose the (non-smooth, constrained) convex problem

$$\min_{\mathbf{a}\in\mathcal{A}} \left[\left\| A\mathbf{a} - d_{\mathrm{obs}} \right\|_{\mathcal{D}}^{2} + \delta_{\mathcal{A}_{+}}(\mathbf{a}) + \lambda \left\| \left\| \mathbf{a}_{\mathsf{r}} \right\|_{\mathrm{L}^{2}(\mathbb{R}_{+})} \right\|_{\mathrm{L}^{1}(\mathbb{R}^{2})} \right]$$

with $\mathbf{r} = (x, y)$.

- How do we solve this optimization problem? Can it be solved?
- Three terms, two non-smooth (with known prox), one smooth (with non-trivial but manageable gradient). Convex problem, but existance and unicity not given (function spaces).
- Do we need forward-backward primal-dual splitting?

We have $d_{obs} \in D_+$ and want to recover $a \in A_+$. We propose the (non-smooth, constrained) convex problem

$$\min_{\mathbf{h}\in\mathcal{A}} \left[\|\mathbf{A}\mathbf{a} - \mathbf{d}_{\mathrm{obs}}\|_{\mathcal{D}}^{2} + \delta_{\mathcal{A}_{+}}(\mathbf{a}) + \lambda \left\| \|\mathbf{a}_{\mathsf{r}}\|_{\mathrm{L}^{2}(\mathbb{R}_{+})} \right\|_{\mathrm{L}^{1}(\mathbb{R}^{2})} \right]$$

with $\mathbf{r} = (x, y)$.

- How do we solve this optimization problem? Can it be solved?
- Three terms, two non-smooth (with known prox), one smooth (with non-trivial but manageable gradient). Convex problem, but existance and unicity not given (function spaces).
- Do we need forward-backward primal-dual splitting? No. Not if we can find the prox of the sum of the two non-smooth terms. It is faster (Pustelnik and Condat, 2017).

We have $d_{obs} \in D_+$ and want to recover $a \in A_+$. We propose the (non-smooth, constrained) convex problem

$$\min_{\mathbf{a}\in\mathcal{A}} \left[\left\| A\mathbf{a} - d_{\mathrm{obs}} \right\|_{\mathcal{D}}^{2} + \delta_{\mathcal{A}_{+}}(\mathbf{a}) + \lambda \left\| \left\| \mathbf{a}_{\mathsf{r}} \right\|_{\mathrm{L}^{2}(\mathbb{R}_{+})} \right\|_{\mathrm{L}^{1}(\mathbb{R}^{2})} \right]$$

with $\mathbf{r} = (x, y)$.

- Do we need forward-backward primal-dual splitting? No. Not if we can find the prox of the sum of the two non-smooth terms. It is faster (Pustelnik and Condat, 2017).
- We showed that the prox of the non-negative group-sparsity regularizer is

$$p_{\mathsf{r}} = \left[a_{\mathsf{r}}\right]_{+} \left(1 - rac{\gamma \lambda}{\left\|\left[a_{\mathsf{r}}
ight]_{+} \right\|_{\mathrm{L}^{2}\left(\left[0,\sigma_{\max}
ight]
ight)}}
ight)_{+}$$

Functional Inverse Diffusion - APG algorithm (Optimization II)

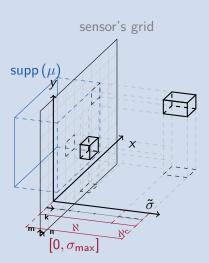
Require: Initial $a^{(0)} \in A_+$, image observation $d_{obs} \in D_+$ **Ensure:** A solution $a_{opt} \in A_+$

1:
$$b^{(0)} \leftarrow a^{(0)}, i \leftarrow 0$$

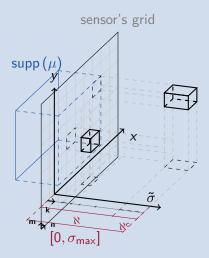
2: **repeat**
3: $i \leftarrow i + 1, \alpha \leftarrow \frac{t(i-1)-1}{t(i)}$
4: $a^{(i)} \leftarrow b^{(i-1)} - \sigma_{\max}^{-1} A^* \left(Ab^{(i-1)} - d_{obs}\right)$
5: **for all r** $\in \mathbb{R}^2$ **do**
6: $a_{\mathbf{r}}^{(i)} \leftarrow \left[a_{\mathbf{r}}^{(i)}\right]_+ \left(1 - \frac{(2\sigma_{\max})^{-1}\lambda}{\left\|\left[a_{\mathbf{r}}^{(i)}\right]_+\right\|_{L^2([0,\sigma_{\max}])}}\right)$
7: **end for**
8: $b^{(i)} \leftarrow a^{(i)} + \alpha \left(a^{(i)} - a^{(i-1)}\right)$

9: **until** convergence 10: $a_{\text{opt}} \leftarrow a^{(i)}$

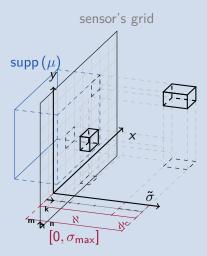
Sequences of t(i) can be chosen as (Bech and Teboulle, 2009) or as (Chambolle and Dossal, 2015).



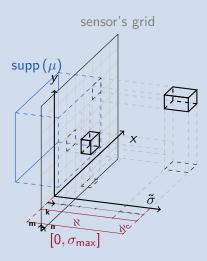
 Spatial grid given by camera sensor



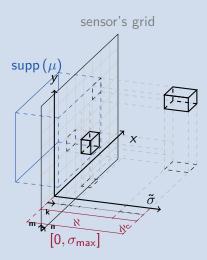
- Spatial grid given by camera sensor
- σ-grid with different levels of detail



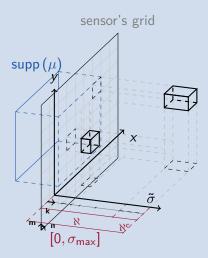
- Spatial grid given by camera sensor
- σ-grid with different levels of detail
- Inner approximation paradigm (step-constant functions)



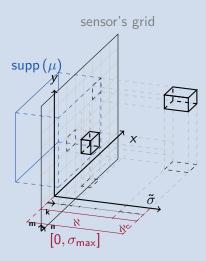
- Spatial grid given by camera sensor
- σ-grid with different levels of detail
- Inner approximation paradigm (step-constant functions)
- Choice of normalization in restriction and extension operators



- Spatial grid given by camera sensor
- σ-grid with different levels of detail
- Inner approximation paradigm (step-constant functions)
- Choice of normalization in restriction and extension operators
- Resulting algorithm can be reasoned as discrete APG



- Spatial grid given by camera sensor
- σ-grid with different levels of detail
- Inner approximation paradigm (step-constant functions)
- Choice of normalization in restriction and extension operators
- Resulting algorithm can be reasoned as discrete APG
- The typical size of the variable a[m, n, k] to recover will be 2048² × 6 = 25 · 10⁶



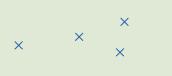
- Spatial grid given by camera sensor
- σ-grid with different levels of detail
- Inner approximation paradigm (step-constant functions)
- Choice of normalization in restriction and extension operators
- Resulting algorithm can be reasoned as discrete APG
- The typical size of the variable a[m, n, k] to recover will be 2048² × 6 = 25 · 10⁶
- Different kernel approximations are considered

Evaluation on Synthetic Data

Besides thorough human testing on real data, we can evaluate our approach on synthetic data. To evaluate the location accuracy, we run 10000 iterations of the algorithm, find spatial maxima and threshold them optimally, and, defining a tolerance of $\Delta = 3 \text{ pix}$ we compute the detection metrics

$$\mathrm{pre} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}\text{, }\mathrm{rec} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}\text{, and }\mathrm{F1} = \frac{2\,\mathrm{pre}\cdot\mathrm{rec}}{\mathrm{pre} + \mathrm{rec}}$$

Example



×: Real cells

Evaluation on Synthetic Data

Besides thorough human testing on real data, we can evaluate our approach on synthetic data. To evaluate the location accuracy, we run 10000 iterations of the algorithm, find spatial maxima and threshold them optimally, and, defining a tolerance of $\Delta = 3 \text{ pix}$ we compute the detection metrics

$$\mathrm{pre} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}\text{, } \mathrm{rec} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}\text{, and } \mathrm{F1} = \frac{2\,\mathrm{pre}\cdot\mathrm{rec}}{\mathrm{pre} + \mathrm{rec}}$$

Example

$$+$$
 $+$ $+$ $+$ \times \times

×: Real cells +:

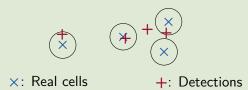
+: Detections

Evaluation on Synthetic Data

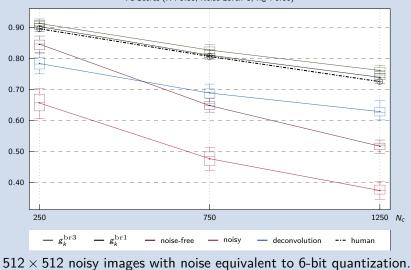
Besides thorough human testing on real data, we can evaluate our approach on synthetic data. To evaluate the location accuracy, we run 10000 iterations of the algorithm, find spatial maxima and threshold them optimally, and, defining a tolerance of $\Delta = 3 \text{ pix}$ we compute the detection metrics

$$pre = \frac{TP}{TP + FP}$$
, $rec = \frac{TP}{TP + FN}$, and $F1 = \frac{2 pre \cdot rec}{pre + rec}$

Example

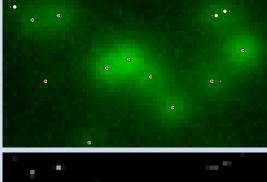


Results on Synthetic Data (I)



F1-Scores (λ : 0.50, Noise Level: 3, λ_d : 0.00)

Results on Synthetic Data (II)

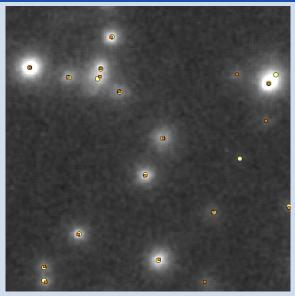


True positions (orange triangles) and detections (yellow circles).

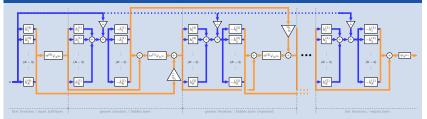
.

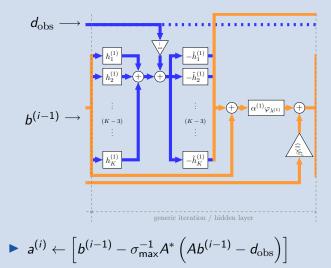
Pixels' contr. to the regularizer, i.e., $\sqrt{\int a^2(x, y, \sigma) d\sigma}$.

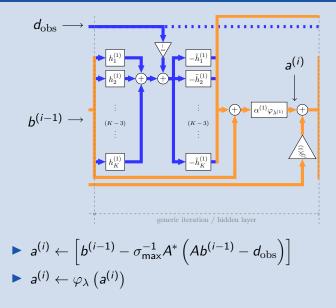
Results on Real Data

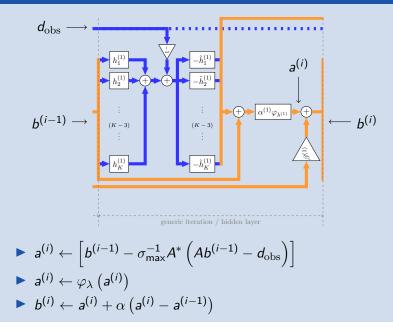


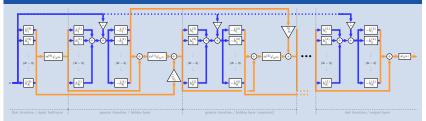
Detection results (yellow circles) and human labeling (orange squares). F1-Score relative to human, 0.9 (whole image).











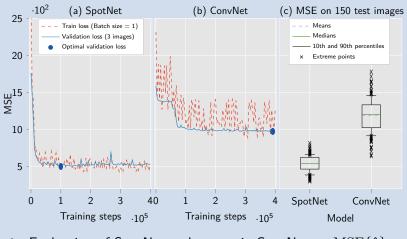
$$a^{(i)} \leftarrow \left[b^{(i-1)} - \sigma_{\max}^{-1} A^* \left(A b^{(i-1)} - d_{obs} \right) \right]$$

$$a^{(i)} \leftarrow \varphi_{\lambda} \left(a^{(i)} \right)$$

$$b^{(i)} \leftarrow a^{(i)} + \alpha \left(a^{(i)} - a^{(i-1)} \right)$$

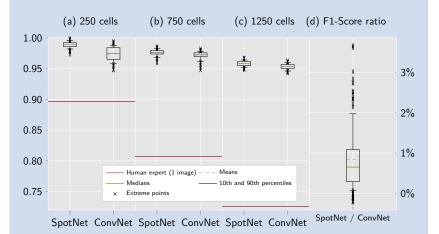
Based on the learned gradient descent of (Gregor and LeCun, 2010), recently explored by (Giryes, Eldar et al., 2018).

Results for SpotNet with L = 3 and smaller kernels



Evaluation of SpotNet and a generic ConvNet on MSE{â}.
 Trained on 7 images with 1250 cells.

Results for SpotNet with L = 3 and smaller kernels



- Evaluation of SpotNet and a generic ConvNet on F1 score as above.
- Trained on 7 images with 1250 cells.



Thank you

Please, feel free to ask questions.

January 14, 2019 at 🌑 CosmoStat

