Supplemental Material -
Single Image Object Counting and Localizing using Active-Learning

Inbar Huberman-Spiegelglas Raanan Fattal
{inbar.huberman1, raanan.fattal}@mail.huji.ac.il
School of Computer Science and Engineering
The Hebrew University of Jerusalem, Israel

Appendix

Method Pseudo-Code

Algorithm 1 summarizes all the steps of our active-learning process.

Algorithm 1: Method Pseudo-Code.

```
Input  : Input image $I$, user-marked bounding window $B$
Output: detected repeating object coordinates $\mathcal{O}$
/* initialization                                          */
$ncc = NCC(I, B)$
$\mathcal{P} = MaxSup(ncc \geq 0.85), \mathcal{N} = ncc \leq 0$
train $CNN$ on $\mathcal{P}, \mathcal{N}$
$C = CNN(I)$
while not terminated do
    $C^s(\mathbf{x}) = MaxSup(C(\mathbf{x}))$
    /* Extract potential locations                          */
    $\mathcal{W} = \{\mathbf{x} | C^s(\mathbf{x}) > 0\}$
    /* Associate potential windows with
       labeled coordinate                                   */
    $l_w, d_w = GetNearestLabel(\mathcal{W}, \mathcal{P}, \mathcal{N})$
    $W^{\mathcal{P}} = \{\mathbf{x} \in \mathcal{W} | l_w = Pos.\}$, $W^{\mathcal{N}} = \{\mathbf{x} \in \mathcal{W} | l_w = Neg.\}$
    /* Clustering each set                                  */
    $\Theta^{\mathcal{P}} = Kmeans(W^{\mathcal{P}}, k=10)$, $\Theta^{\mathcal{N}} = Kmeans(W^{\mathcal{N}}, k=10)$
    /* find most distant windows                            */
    $q_i^{\mathcal{P}} = GetTop5Clust(\Theta_i^{\mathcal{P}}, d_w)$,
     $q_i^{\mathcal{N}} = GetTop5Clust(\Theta_i^{\mathcal{N}}, d_w)$
    /* User corrections                                     */
    $\mathcal{L}^{\mathcal{P}}, \mathcal{L}^{\mathcal{N}} = GetUserInput(q_i^{\mathcal{P}}, q_i^{\mathcal{N}})$
    /* Updating label sets                                  */
    $\mathcal{P} = \mathcal{P} \cup \mathcal{L}^{\mathcal{P}}$, $\mathcal{N} = \mathcal{N} \cup \mathcal{L}^{\mathcal{N}}$
    /* further training                                     */
    train $CNN$ on $\mathcal{P}, \mathcal{N}$
    $C = CNN(I)$
$C^s(\mathbf{x}) = MaxSup(C(\mathbf{x}))$
$\mathcal{O} = \{\mathbf{x} | C^s(\mathbf{x}) > 0\}$
```

Hyper-parameter Search and Ablation Study

In order to search the optimal hyper-parameters as well as to evaluate the contribution of the proposed method components, we set up an "automated" version of our method. In this mode, we use a ground-truth, per-pixel, image labeling in order to provide an automated user feedback, as well as an initial bounding window pointing out the object of interest. This allows us to perform extensive tests over multiple images. We used ten images to perform this hyper-parameter search and ablation study, non of which appears in the test sets that we report in Table 1 in the paper.

configuration		NCC		sub-space loss							random querying	cross-entropy
		threshold		m		α						
param. values		0.8	0.9	0.5N	1.5N	0	0.5	1.5	2	2.5		
diff. cnt. err. [%]		+0.62	+0.69	+0.97	+0.91	+0.95	+1.07	+0.62	+0.65	+0.45	+0.77	+0.45

Table 1: Hyper-parameters Search and Ablation Study. Difference in error percentage are reported with respect to changes from the default values used in our method, namely, $\alpha = 1, m = N$ and NCC threshold of 0.85.

The evaluation of these hyper-parameters is summarized in Table 1 which reports the differences of average counting error percentage when the default values are changed to the ones in the table. The parameters explored are: the NCC threshold at the network initialization step in Section 3.1 (the default value is 0.85), the sub-space separation dimension m from Eq. 3 (default value N), as well as its loss weight α suggested in Section 3.4 (default value 1).

Next we evaluate the benefit obtained by the novel components of our method. By setting $\alpha = 0$ we measure the contribution of the sub-space separation loss which, according to Table 1, reduces the average counting error by 15.7%, from 6.95% to 6% (our method's performance in the automated mode). The cluster-based query extraction was compared to a random selection of queries from $\mathcal{W}^{\mathcal{P}}$and $\mathcal{W}^{\mathcal{N}}$. This test shows a reduction of 12.9%, from 6.78% to 6%, in average counting error. Finally, the use of MSE loss in Eq. 2 instead of a cross-entropy loss reduces the counting error by 6.4%, from 7.4% to 6.95% (when using $\alpha = 0$).

We also evaluated the number of user corrections when presenting all the queries along with a positive tentative labels (following the fact that $C(\mathbf{x}) > 0, \forall \mathbf{x} \in \mathcal{W}$), compared to the labels derived from the association with $\mathcal{W}^{\mathcal{P}}$or $\mathcal{W}^{\mathcal{N}}$that we use. This resulted in a reduction of 34% in the average number of user mouse clicks, from 18.5 to 13.8. Note that since both cases consist of the same query extraction scheme, there is no change in the counting error.

Comprehensive Results for User-Study

Below is the full table comparing Artera *et al*. [1], Huberman and Fattal [3] and our method.

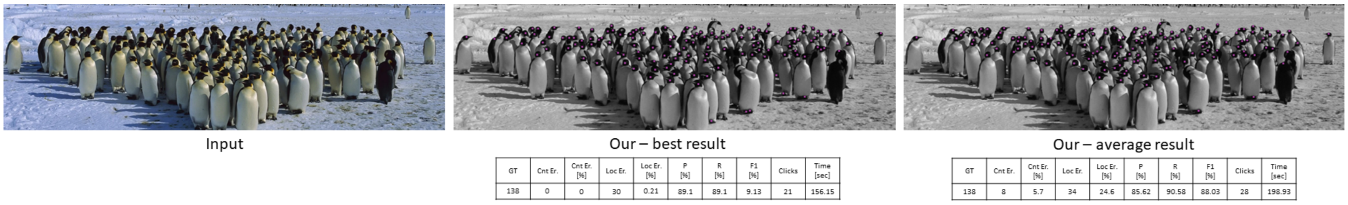
	Artera et al. [1]									Huberman and Fattal [3]									Ours								
image	**Cnt. Er.**	**Cnt. Er.**	**Loc. Er.**	**Loc. Er.**	**P**	**R**	**F1**	**Time**	**Clicks**	**Cnt. Er.**	**Cnt. Er.**	**Loc. Er.**	**Loc. Er.**	**P**	**R**	**F1**	**Time**	**Clicks**	**Cnt. Er.**	**Cnt. Er.**	**Loc. Er.**	**Loc. Er.**	**P**	**R**	**F1**	**Time**	**Clicks**
		[%]		[%]	[%]	[%]	[%]	[sec]			[%]		[%]	[%]	[%]	[%]	[sec]			[%]		[%]	[%]	[%]	[%]	[sec]	
Antarctica	34.4	24.9	102.2	74.1	70.0	43.7	51.9	177.4	14.6	–	–	–	–	–	–	–	–	–	10.1	7.3	37.8	27.3	87.0	86.0	86.6	210.4	22.0
Beach	21.5	7.8	92.0	33.3	80.7	74.6	77.1	186.5	20.3	13.0	4.7	50.3	18.2	93.3	88.2	90.6	263.8	25.0	15.6	5.7	33.6	12.2	92.9	95.4	94.0	178.0	18.0
Beer	18.1	8.7	47.0	22.6	96.5	81.7	88.1	174.3	19.7	1.0	0.5	11.0	5.3	97.6	97.1	97.3	53.5	3.0	1.8	0.9	2.6	1.3	99.0	99.8	99.4	96.3	10.4
Bees	13.9	19.3	53.5	74.3	56.0	55.2	54.1	261.0	16.5	–	–	–	–	–	–	–	–	–	6.0	8.3	27.3	38.0	83.1	78.2	80.4	225.7	16.0
Birds	21.2	9.2	53.4	23.2	97.8	78.5	87.0	206.4	24.8	66.7	29.0	109.3	47.5	81.0	75.1	76.0	290.7	34.0	25.2	10.9	43.8	19.1	92.6	88.9	90.4	205.7	20.8
Birds002	87.1	11.6	287.4	38.3	94.9	65.3	76.7	267.4	29.2	217.0	28.9	253.0	33.7	96.6	68.7	80.3	319.7	36.0	65.6	10.9	123.2	16.4	95.8	87.4	91.4	281.5	18.4
Candles	4.4	5.6	47.0	60.3	91.4	43.8	59.1	165.2	22.8	5.5	7.1	15.5	19.9	21.4	19.9	20.6	72.7	9.0	4.0	5.1	6.0	7.7	98.3	94.0	96.1	87.8	11.5
Cars	12.7	12.9	39.5	40.3	93.8	59.6	72.2	148.9	18.6	18.7	19.0	51.3	52.4	72.4	79.3	75.2	124.3	35.7	3.3	3.3	5.8	5.9	96.6	97.7	97.1	123.6	18.8
CarsBg	233.9	26.7	529.0	60.3	94.1	42.0	57.3	288.0	27.7	30.7	3.5	96.3	11.0	93.4	95.9	94.6	371.6	25.7	37.7	4.3	52.3	6.0	95.3	99.1	97.1	191.5	22.5
CellLrg	26.0	10.9	31.8	13.4	99.2	87.2	92.7	209.0	18.4	1.0	0.4	3.0	0.4	99.2	99.6	99.4	61.9	9.5	0.2	0.1	0.2	0.1	99.9	100.0	100.0	79.0	8.8
CellSml	7.3	3.1	28.0	11.9	99.6	90.7	94.9	163.0	19.5	1.0	0.4	3.0	1.3	99.2	99.6	99.4	85.2	6.0	0.8	0.3	0.8	0.3	99.7	100.0	99.8	100.8	9.4
Chairs	26.7	4.6	113.4	19.5	99.6	81.2	89.0	192.4	28.6	87.0	14.9	149.0	25.6	93.6	79.7	86.0	264.9	27.7	3.4	0.6	4.6	0.8	99.4	99.8	99.6	120.8	12.0
CokeDiet	3.3	8.0	32.6	79.5	68.1	41.5	50.9	180.0	12.0	–	–	–	–	–	–	–	–	–	2.6	6.3	3.8	9.3	98.0	92.7	95.2	112.3	7.8
CokeReg	2.9	8.8	14.2	43.0	92.6	62.4	74.4	160.8	12.0	–	–	–	–	–	–	–	–	–	0.8	2.4	0.8	2.4	97.8	100.0	98.9	77.7	3.4
Cookies	10.0	5.8	97.3	56.2	85.4	52.2	64.0	157.8	20.0	7.5	4.3	16.0	9.2	98.5	92.2	95.2	83.0	26.5	3.2	1.8	8.4	4.9	98.5	96.6	97.5	103.4	13.4
Crabs	8.3	4.3	16.8	8.7	99.6	89.5	94.2	210.3	14.5	53.8	27.8	144.3	74.7	67.1	62.7	62.0	231.9	37.3	4.6	2.4	5.8	3.0	97.4	99.7	98.5	131.8	15.2
Crowd	26.9	6.7	216.8	53.7	87.1	54.1	66.5	183.0	26.6	67.0	16.6	285.0	70.5	67.7	56.4	61.5	233.4	37.0	32.8	8.1	136.8	33.9	85.6	79.8	82.5	172.3	20.6
Discussion	18.4	24.5	84.2	112.3	43.1	36.6	39.3	187.1	16.9	–	–	–	–	–	–	–	–	–	6.7	8.9	40.0	53.3	73.1	73.3	73.0	262.3	21.2
Fish097	3.5	6.0	17.3	29.9	92.2	77.0	83.8	116.0	13.7	6.0	10.3	15.0	25.9	90.6	83.6	86.7	70.2	12.0	5.8	10.0	9.8	16.9	87.9	96.6	92.0	137.5	15.8
Fish107	9.6	14.1	14.3	21.0	96.3	82.4	88.7	144.5	16.5	7.0	10.3	11.0	16.2	88.0	97.1	92.3	76.5	8.0	1.8	2.6	3.3	4.8	96.5	98.9	97.7	75.6	9.8
Flowers	22.4	18.2	66.4	54.0	98.7	46.7	63.3	173.0	25.0	22.8	18.5	88.4	71.9	64.9	65.4	64.6	278.1	36.2	24.8	20.2	37.2	30.2	93.8	74.8	83.2	149.7	13.8
Hats	25.7	8.1	115.2	36.3	92.8	69.1	78.9	224.3	19.8	85.2	26.9	256.8	81.0	60.2	62.7	60.1	293.5	31.8	9.0	2.8	53.8	17.0	91.8	91.2	91.5	142.4	18.8
Logs	20.2	9.8	96.9	47.0	99.9	53.1	69.1	154.0	20.9	11.5	5.6	12.5	6.1	95.5	98.5	97.0	142.8	13.0	4.6	2.2	10.2	5.0	98.6	96.4	97.5	107.5	22.4
Matches	13.4	4.9	76.6	28.0	99.1	72.6	83.5	194.6	31.8	1.3	0.5	2.3	0.8	99.8	99.4	99.6	32.0	9.5	0.6	0.2	0.6	0.2	99.8	100.0	99.9	86.7	8.8
Oranges	19.0	11.7	53.4	32.8	90.4	75.2	82.0	197.6	19.6	33.0	20.2	91.0	55.8	77.6	63.2	69.2	292.6	26.8	9.6	5.9	25.6	15.7	92.0	92.5	92.2	108.9	20.6
Parasol	6.8	13.6	29.8	59.5	74.1	64.0	68.4	155.0	12.5	6.0	12.0	8.5	17.0	88.4	97.5	92.5	192.0	9.3	2.8	5.6	5.6	11.2	94.4	94.8	94.5	93.9	12.2
Peas	29.7	24.5	68.4	56.5	100.0	43.5	60.3	153.6	21.2	19.7	16.3	35.7	29.5	80.8	93.4	86.4	174.3	30.0	18.7	15.5	24.1	20.0	97.4	82.3	89.1	123.8	14.3
Pills	6.1	6.7	4.8	5.2	99.1	95.7	97.3	153.8	19.6	1.0	1.1	3.0	3.3	97.8	98.9	98.3	60.0	19.5	0.2	0.2	0.2	0.2	99.8	100.0	99.9	73.0	8.8
RealCells	19.5	6.0	105.4	32.4	97.8	69.2	80.5	139.8	21.8	2.0	0.6	19.0	5.8	97.1	97.1	97.1	95.0	12.5	2.5	0.8	12.5	3.8	96.9	98.2	97.5	175.2	14.4
Sheep	26.0	10.0	134.0	51.5	93.1	52.3	66.8	236.2	24.7	39.3	15.1	150.0	57.7	70.8	77.2	73.3	227.5	31.3	24.8	9.6	67.5	26.0	86.2	89.0	87.4	262.5	21.5
Soldiers	42.1	28.6	107.2	72.9	84.0	33.2	47.4	167.2	24.4	27.0	18.4	85.0	57.8	75.8	61.9	68.2	193.2	23.0	18.6	12.7	68.2	46.4	78.9	73.7	75.6	140.6	22.6
Wall	24.2	14.7	120.2	73.3	81.3	33.4	46.2	155.8	22.4	12.0	7.3	25.5	15.5	90.7	94.2	92.4	93.1	16.0	11.6	7.0	20.8	12.7	92.4	95.5	93.8	100.5	19.2
Water	14.4	10.9	16.6	12.6	97.7	89.4	93.0	174.4	17.2	2.0	1.5	3.0	2.3	99.2	98.5	98.9	51.9	11.0	0.3	0.3	0.3	0.3	99.7	100.0	99.9	95.8	6.2
Average	26.0	11.9	88.3	43.6	89.3	63.5	72.7	183.6	20.4	30.2	11.5	71.2	29.2	84.2	82.2	82.7	168.9	21.5	**10.9**	**5.6**	**26.5**	**13.7**	**93.8**	**92.5**	**93.0**	**140.4**	**15.1**
SD	40.4	7.1	98.9	24.6	13.5	18.5	16.2	38.2	5.1	44.9	9.6	83.4	26.1	17.4	19.1	18.1	101.3	11.2	14.1	4.9	33.6	14.1	6.6	8.5	7.3	58.4	5.5

Table 2: User-Study Results. The columns report (left-to-right) the average counting error (Mean Absolute Error), average counting error percentage, average localization error (false-positives plus false-negatives), average localization error percentage, Precision, Recall, F-score, interactive session time and number of user mouse-clicks. This is repeated for each method. The average and standard deviation of the counting and localization of each method are presented in the two last rows of the table. The rows list the test images used. Images Birds002, Fish097, Fish107, and Bees were taken from the Small Objects dataset [5], CellLrg, CellSml from [4], Soldiers from [6] and RealCells from [2].

Qualitative Results

Below we present the 33 test images along with our outputs produced by the user-study.

- For each image we present the best result and average result obtained in our user study in terms of localization error
- Our result images are shown in gray-scale and our localizations are indicated in purple dots
- For some images - which appear to be easy for our method to obtain accurate results we show the worst result obtained in the user study (this is indicated in the relevant figures)



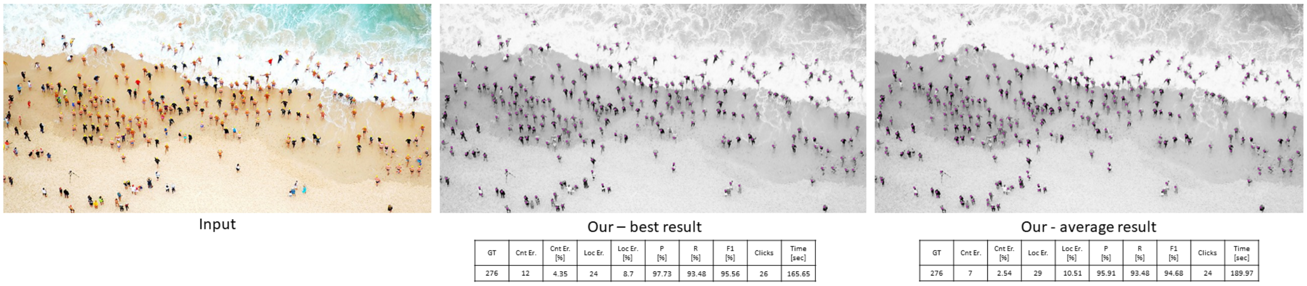
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
138	0	0	30	0.21	89.1	89.1	9.13	21	156.15

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
138	8	5.7	34	24.6	85.62	90.58	88.03	28	198.93

Figure 1: Antarctica.



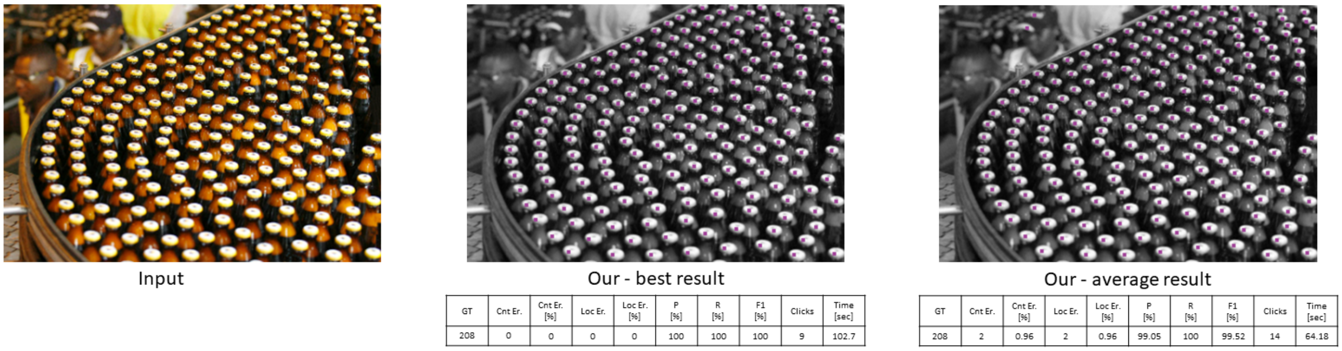
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
276	12	4.35	24	8.7	97.73	93.48	95.56	26	165.65

Our - average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
276	7	2.54	29	10.51	95.91	93.48	94.68	24	189.97

Figure 2: Beach.



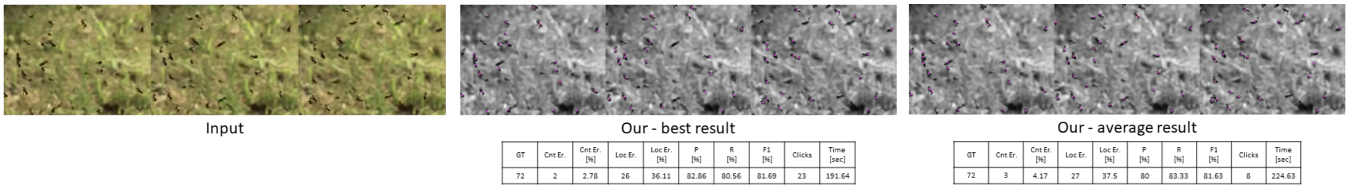
Our - best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
208	0	0	0	0	100	100	100	9	102.7

Our - average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
208	2	0.96	2	0.96	99.05	100	99.52	14	64.18

Figure 3: Beer.



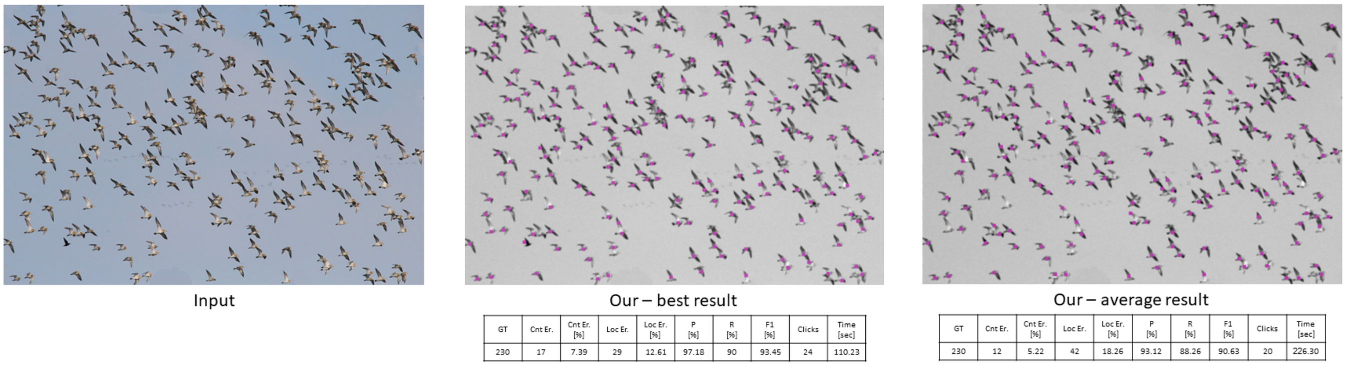
Our - best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
72	2	2.78	26	36.11	82.86	80.56	81.69	23	191.64

Our - average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
72	3	4.17	27	37.5	80	83.33	81.63	8	224.63

Figure 4: Bees. 3 frames, small object dataset [5]



Input

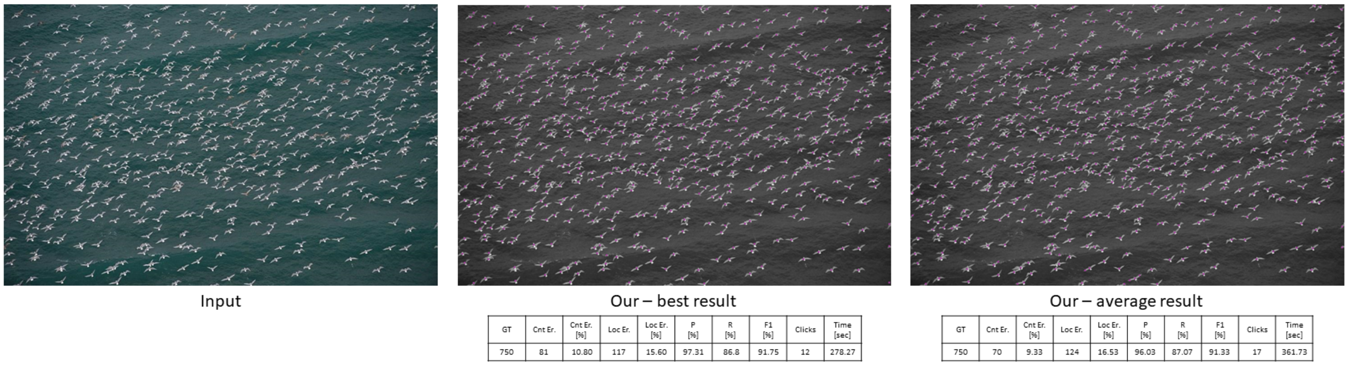
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
230	17	7.39	29	12.61	97.18	90	93.45	24	110.23

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
230	12	5.22	42	18.26	93.12	88.26	90.63	20	226.30

Figure 5: Birds.



Input

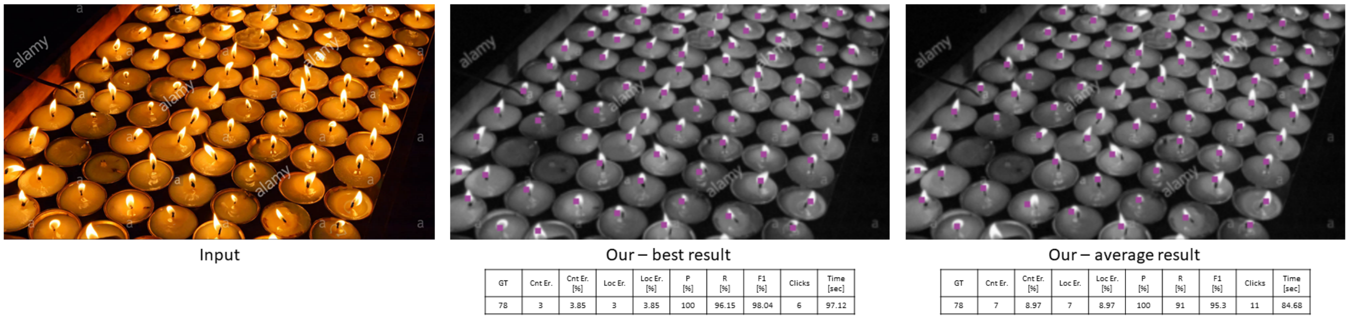
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
750	81	10.80	117	15.60	97.31	86.8	91.75	12	278.27

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
750	70	9.33	124	16.53	96.03	87.07	91.33	17	361.73

Figure 6: Birds002. small object dataset [5]



Input

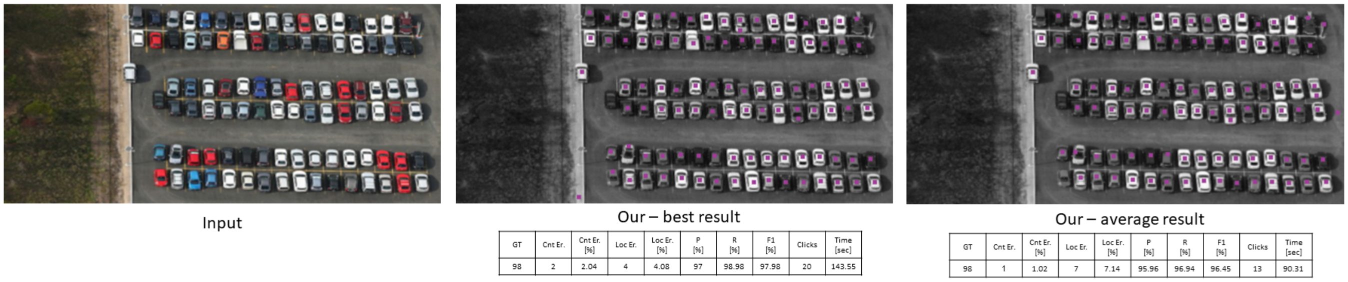
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
78	3	3.85	3	3.85	100	96.15	98.04	6	97.12

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
78	7	8.97	7	8.97	100	91	95.3	11	84.68

Figure 7: Candles.



Input

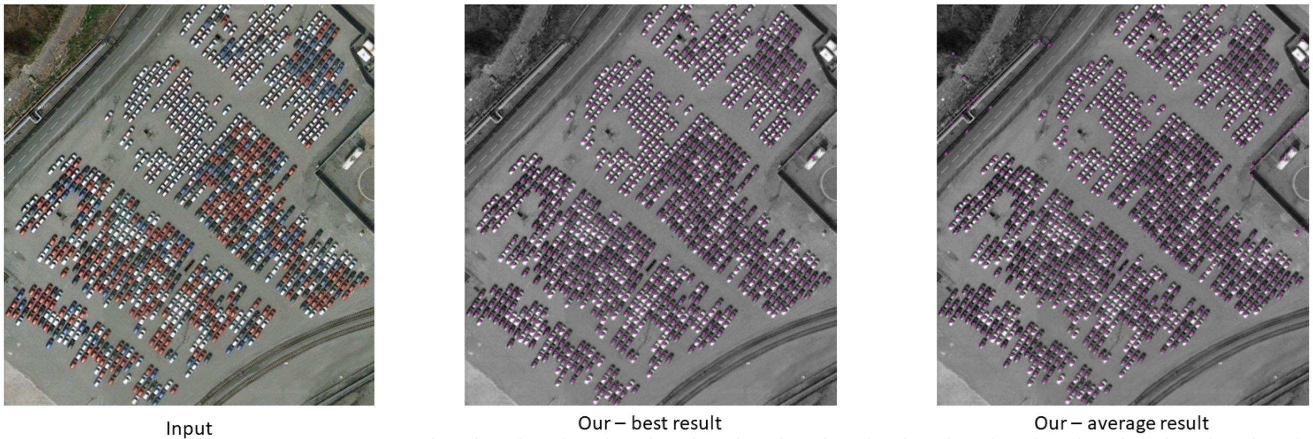
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
98	2	2.04	4	4.08	97	98.98	97.98	20	143.55

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
98	1	1.02	7	7.14	95.96	96.94	96.45	13	90.31

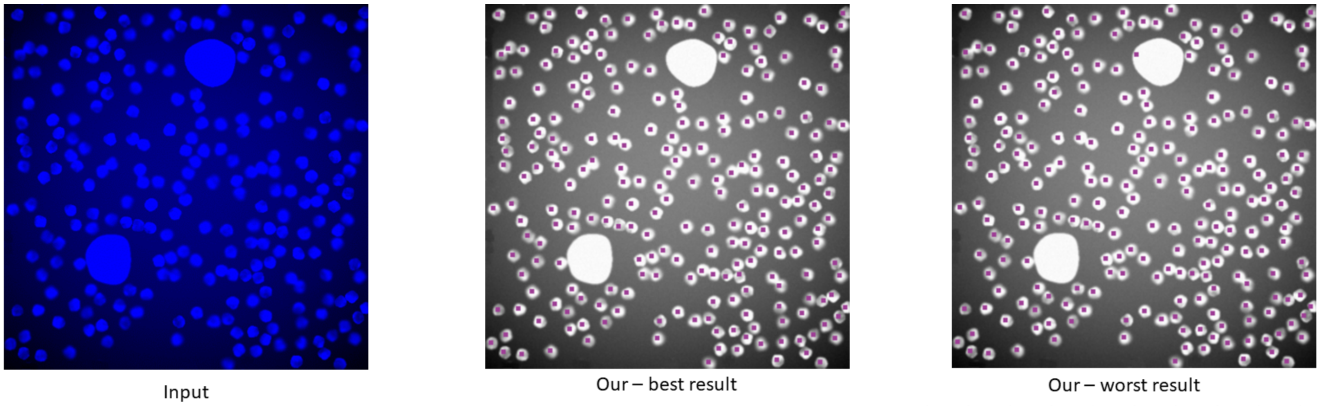
Figure 8: Cars.



GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
877	3	0.34	7	0.80	99.43	99.77	99.60	23	134.97

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
877	49	5.59	73	8.32	93.41	98.63	95.95	23	245.13

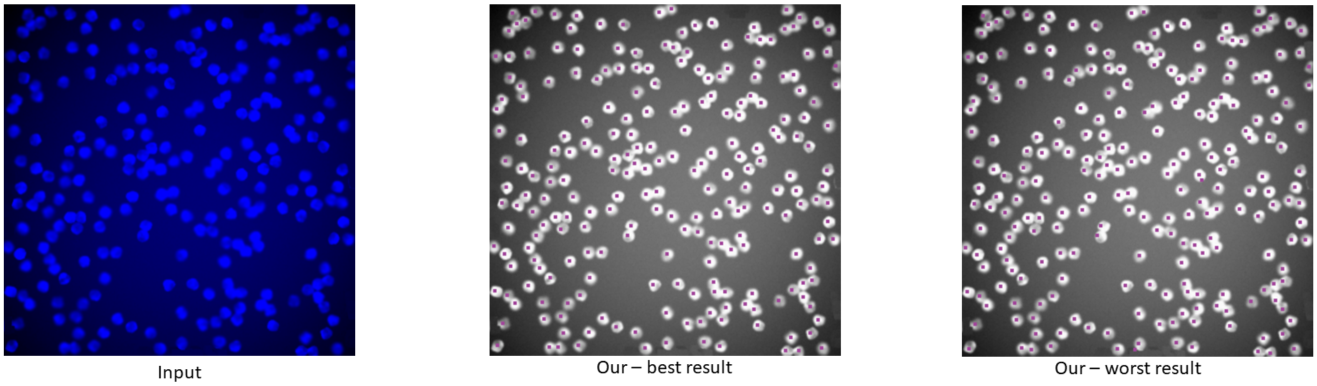
Figure 9: CarsBg.



GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
237	0	0	0	0	100	100	100	12	196.31

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
237	1	0.42	1	0.42	99.58	100	99.79	7	42.79

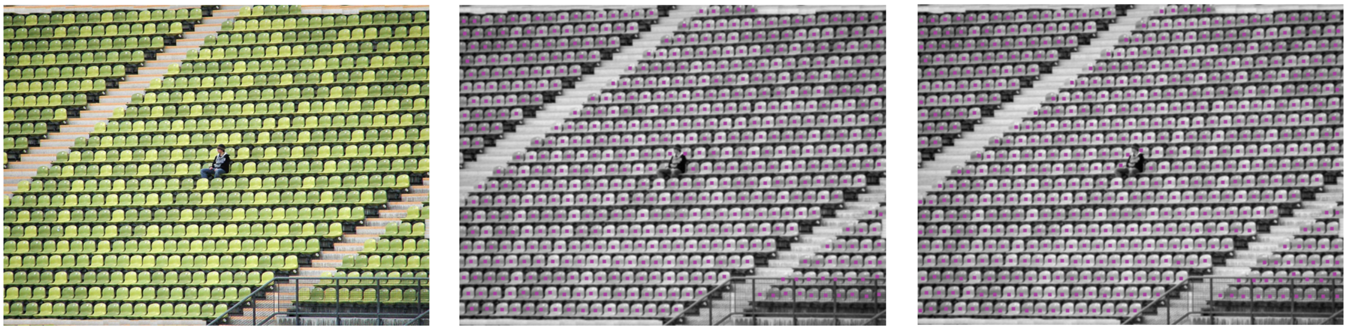
Figure 10: CellLrg. fluorescence microscopy cell images [4].



GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
236	0	0	0	0	100	100	100	8	62.5

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
236	2	0.85	2	0.85	99.16	100	99.58	17	140.56

Figure 11: CellSml. fluorescence microscopy cell images [4].



Input

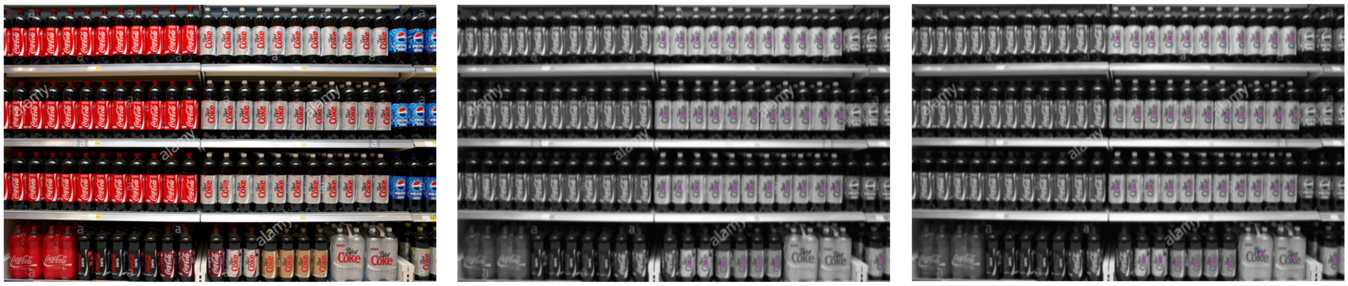
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
583	1	0.17	1	0.17	100	99.83	99.91	12	152.29

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
582	5	0.86	5	0.86	99.15	100	99.57	9	95.12

Figure 12: Chairs.



Input

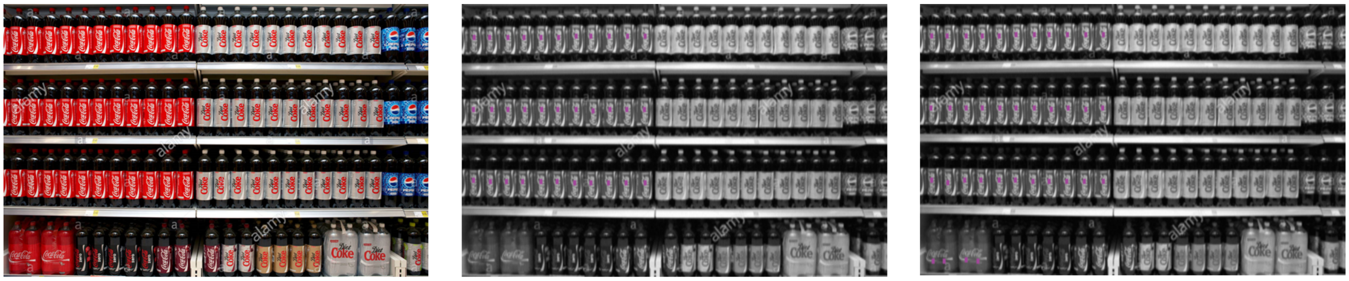
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
40	0	0	0	0	100	100	100	7	103.12

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
40	2	5	2	5	95.24	100	97.56	8	69.31

Figure 13: CokeDiet.



Input

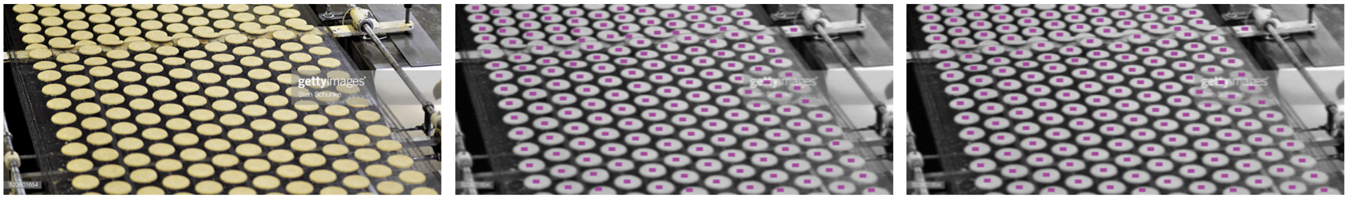
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
33	0	0	0	0	100	100	100	6	85.06

Our – worst result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
33	4	12.12	4	12.12	89.19	100	94.29	3	59.01

Figure 14: CokeReg.



Input

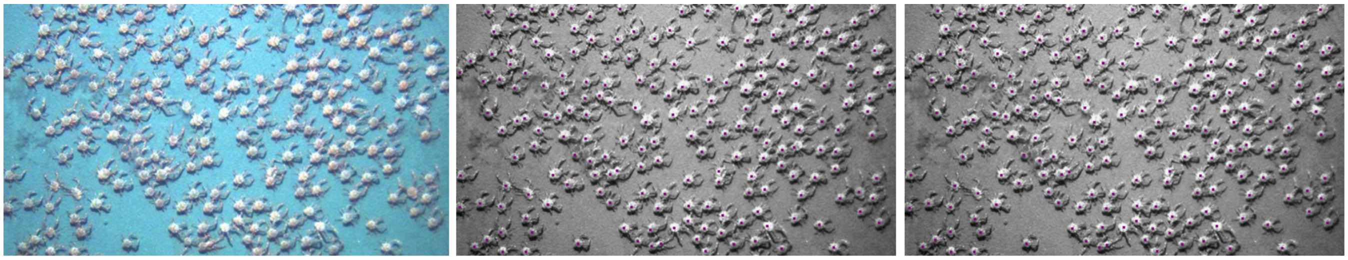
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
174	1	0.005	3	1.72	99.42	98.85	99.13	19	84.3

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
174	3	1.72	6	3.44	98.83	97.12	97.97	12	149.66

Figure 15: Cookies.



Input

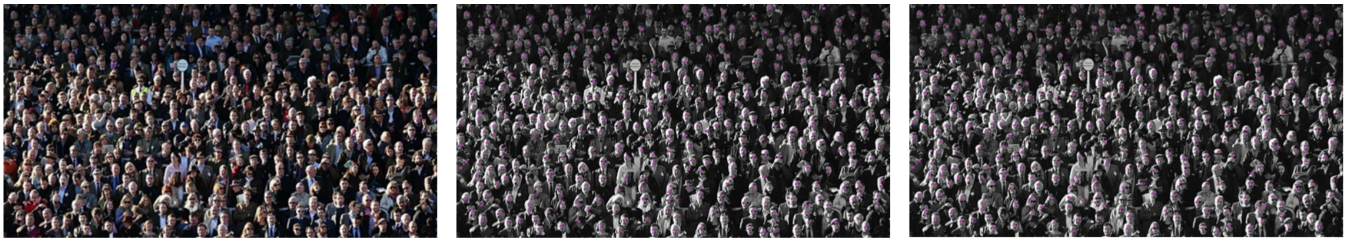
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
193	2	1.04	2	1.04	98.97	100	99.48	10	115.74

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
193	3	1.55	5	2.59	97.96	99.48	98.71	13	93.55

Figure 16: Crabs.



Input

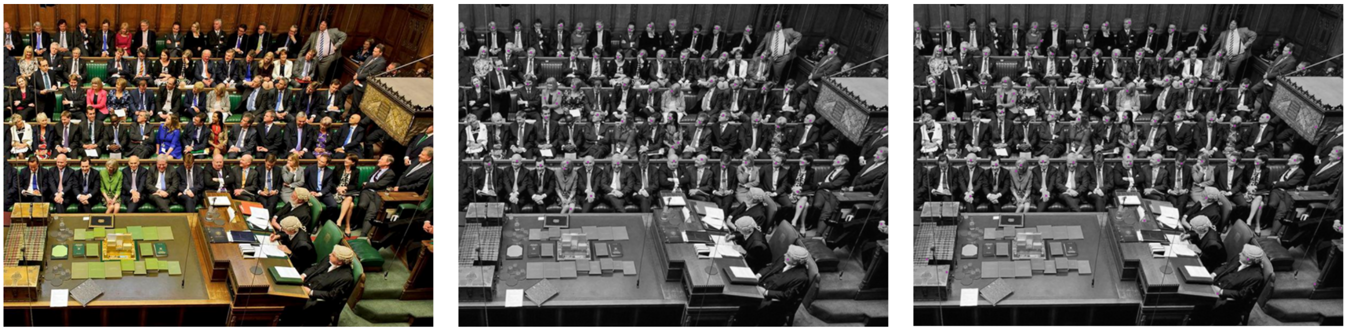
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
404	41	10.15	125	30.94	88.43	79.46	83.70	17	187.46

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
404	47	11.63	139	34.41	87.11	76.98	81.73	20	234.82

Figure 17: Crowd.



Input

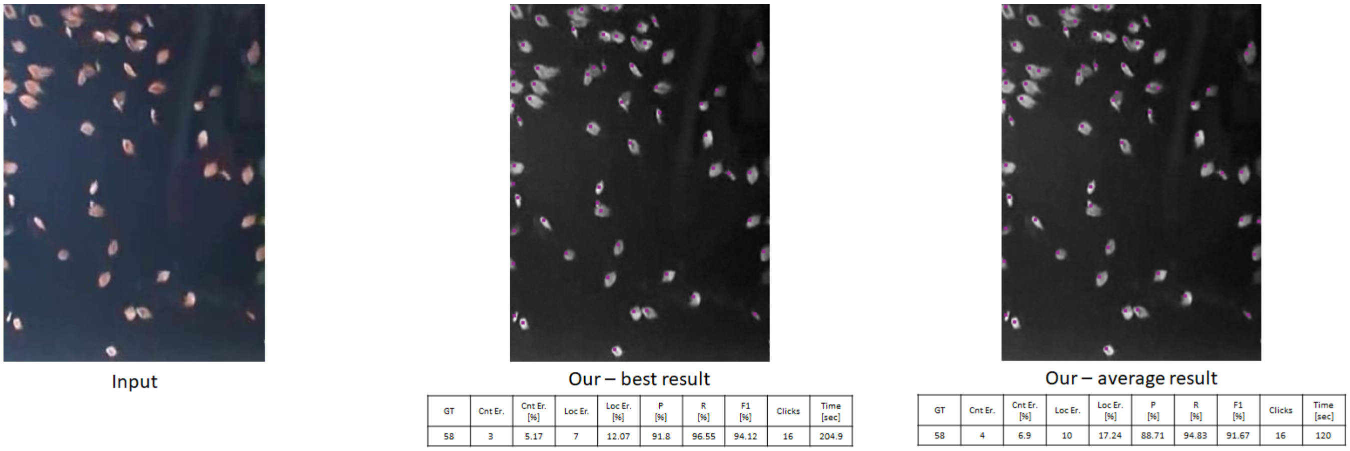
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
75	2	2.67	26	34.67	81.82	84	82.89	19	226.87

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
75	10	13.33	34	45.33	74.12	84	78.75	21	334.48

Figure 18: Discussion.



Input

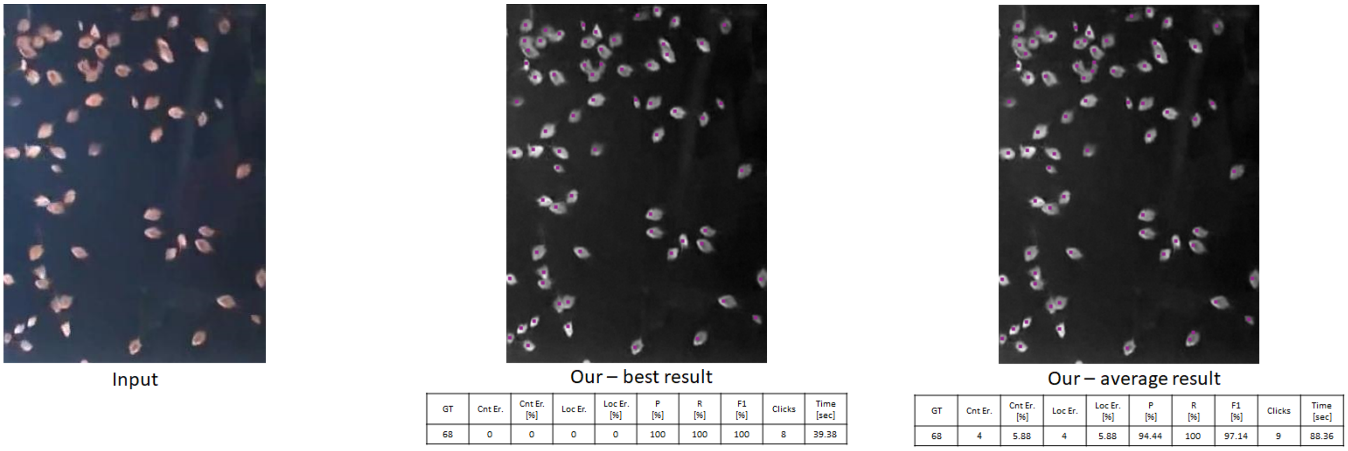
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
58	3	5.17	7	12.07	91.8	96.55	94.12	16	204.9

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
58	4	6.9	10	17.24	88.71	94.83	91.67	16	120

Figure 19: Fish097. small object dataset [5].



Input

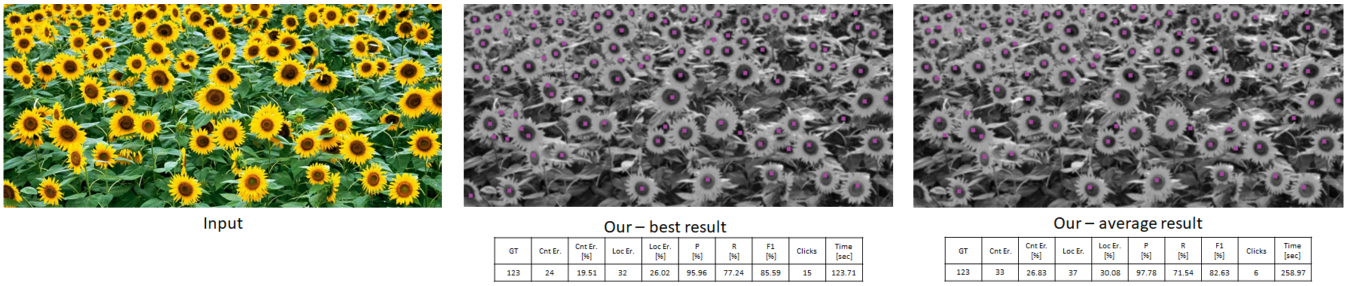
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
68	0	0	0	0	100	100	100	8	39.38

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
68	4	5.88	4	5.88	94.44	100	97.14	9	88.36

Figure 20: Fish107. small object dataset [5].



Input

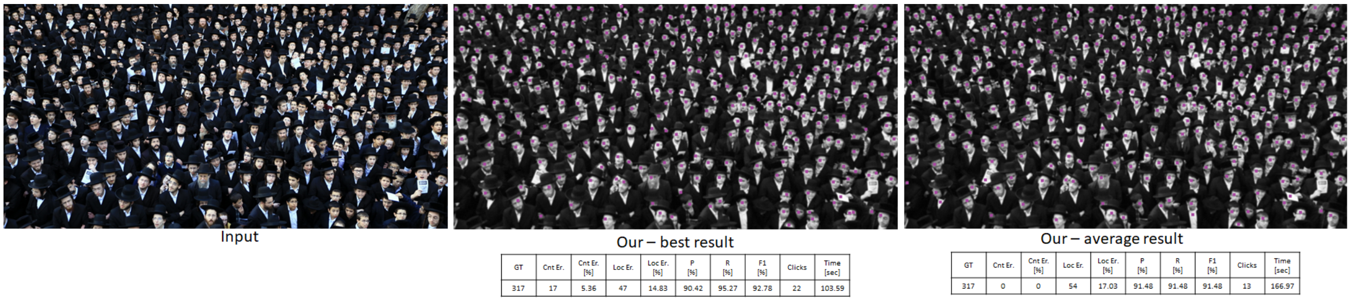
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
123	24	19.51	32	26.02	95.96	77.24	85.59	15	123.71

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
123	33	26.83	37	30.08	97.78	71.54	82.63	6	258.97

Figure 21: Flowers.



Input

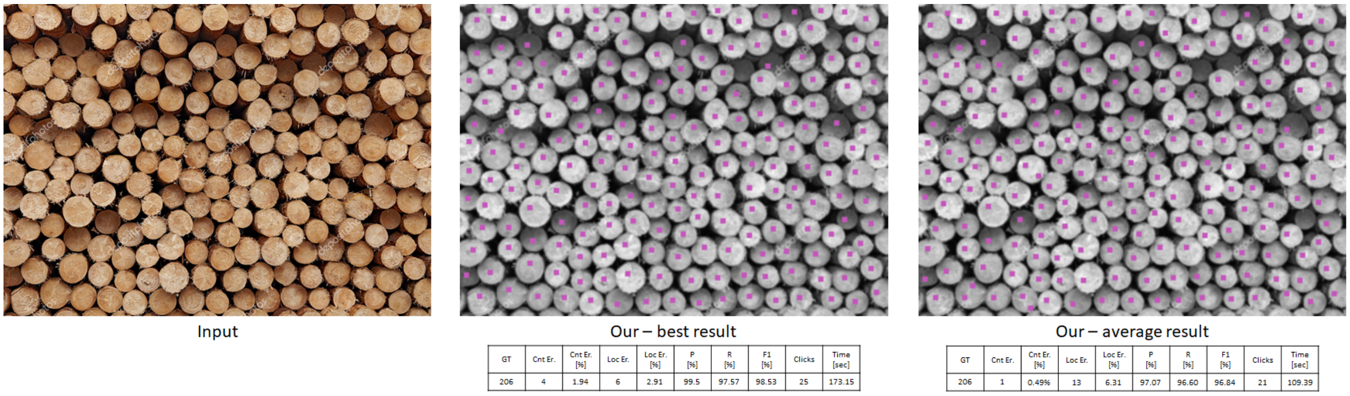
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
317	17	5.36	47	14.83	90.42	95.27	92.78	22	103.59

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
317	0	0	54	17.03	91.48	91.48	91.48	13	166.97

Figure 22: Hats.



Input

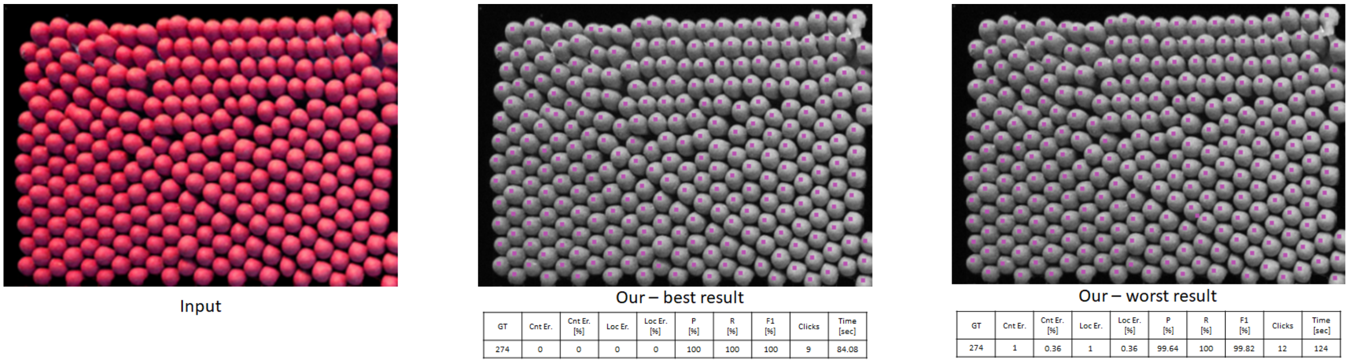
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
206	4	1.94	6	2.91	99.5	97.57	98.53	25	173.15

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
206	1	0.49%	13	6.31	97.07	96.60	96.84	21	109.39

Figure 23: Logs.



Input

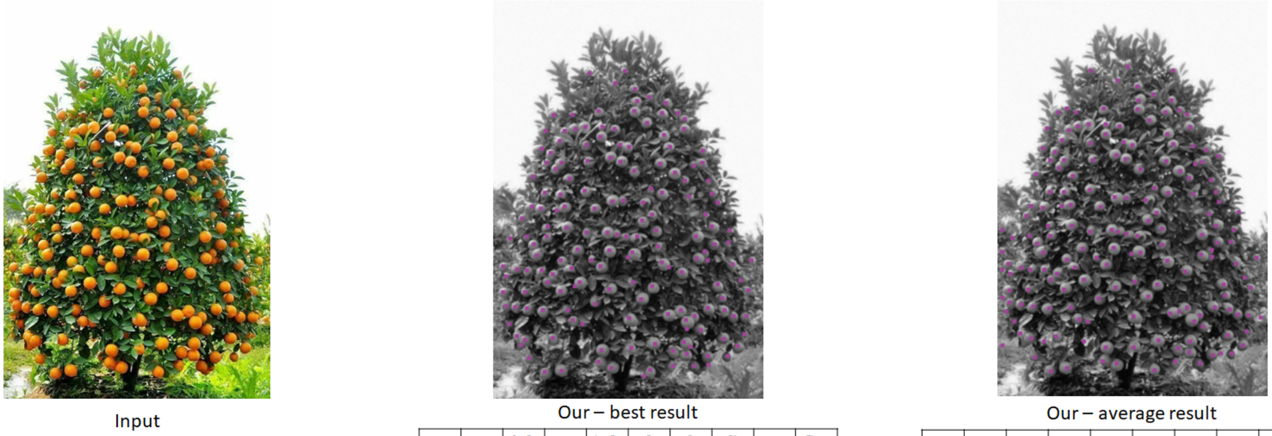
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
274	0	0	0	0	100	100	100	9	84.08

Our – worst result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
274	1	0.36	1	0.36	99.64	100	99.82	12	124

Figure 24: Matches.



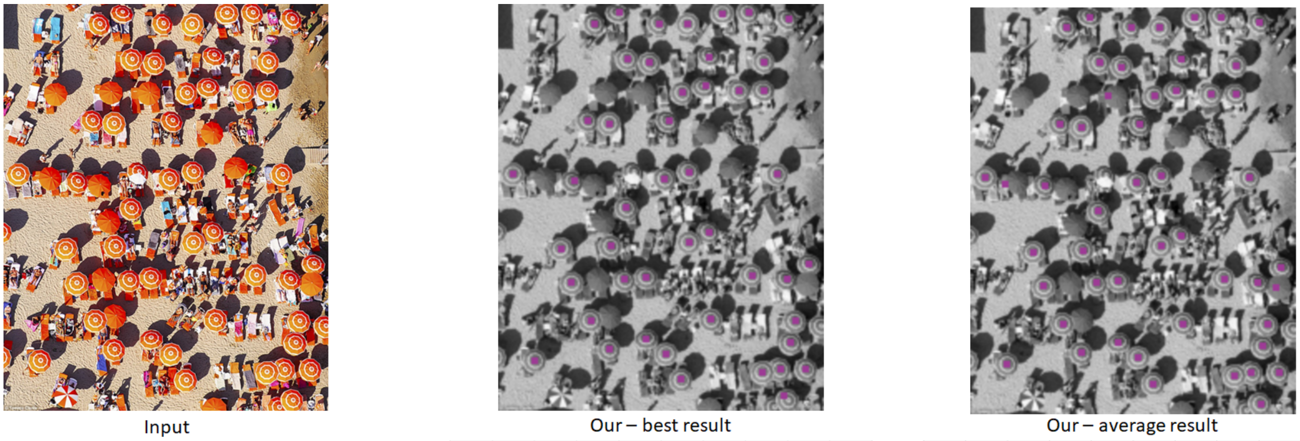
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
163	5	3.07	21	12.88	92.26	95.09	93.66	21	84.5

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
163	14	8.59	26	15.95	88.7	96.32	92.35	23	86.22

Figure 25: Oranges.



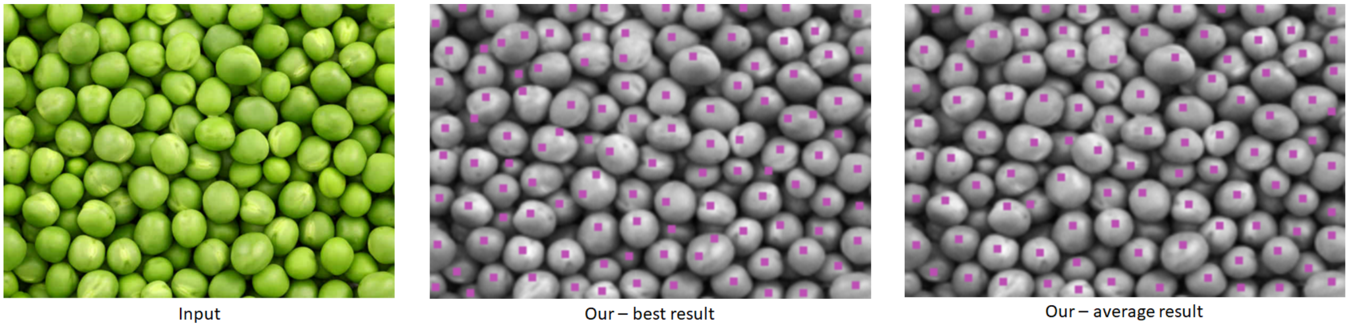
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
50	3	6	3	6	100	94	96.91	13	79.07

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
50	0	0	6	12	94	94	94	8	96.115

Figure 26: Parasol.



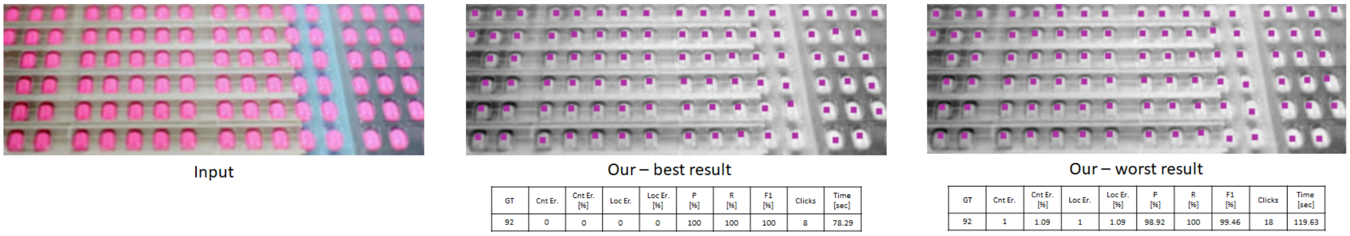
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
121	6	4.96	18	14.88	94.78	90.08	92.37	17	99.68

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
121	20	16.53	24	19.83	98.02	81.82	89.19	13	144.40

Figure 27: Peas.



Input

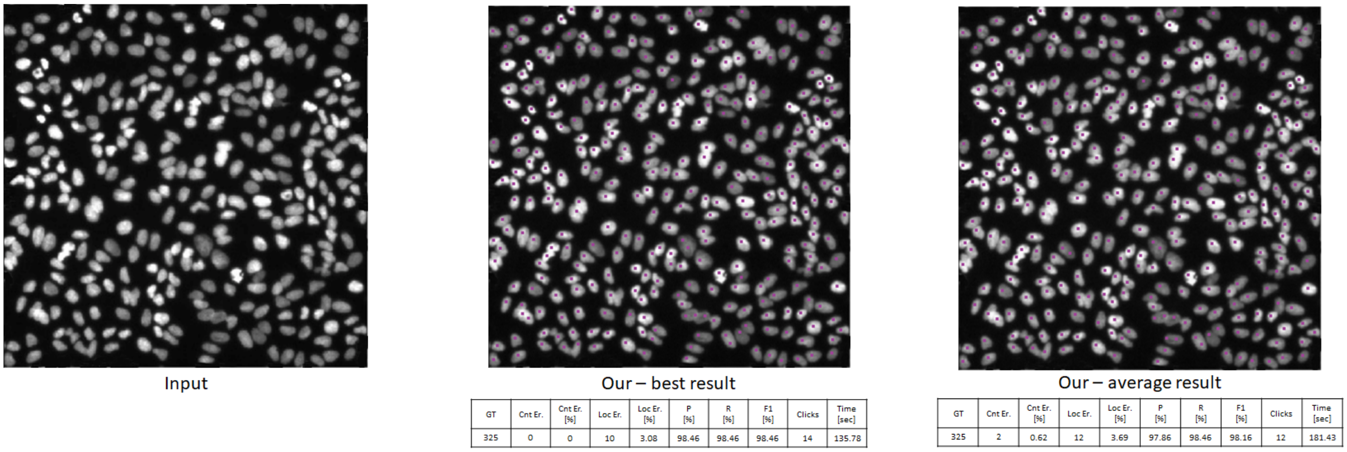
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
92	0	0	0	0	100	100	100	8	78.29

Our – worst result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
92	1	1.09	1	1.09	98.92	100	99.46	18	119.63

Figure 28: Pills.



Input

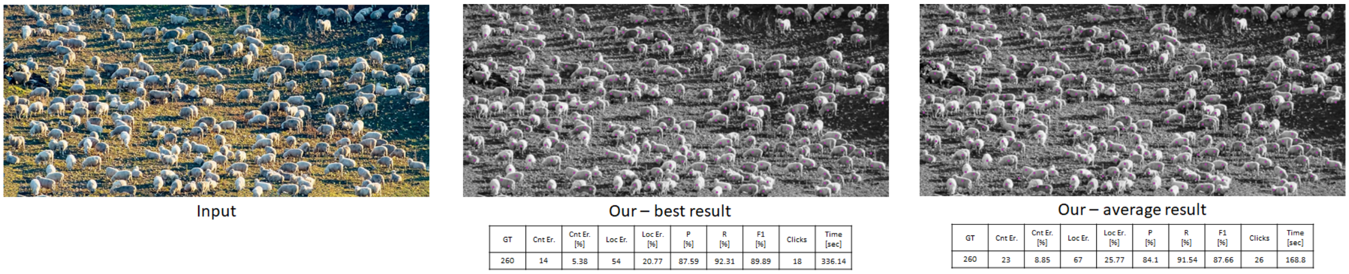
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
325	0	0	10	3.08	98.46	98.46	98.46	14	135.78

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
325	2	0.62	12	3.69	97.86	98.46	98.16	12	181.43

Figure 29: RealCells. Taken from [2].



Input

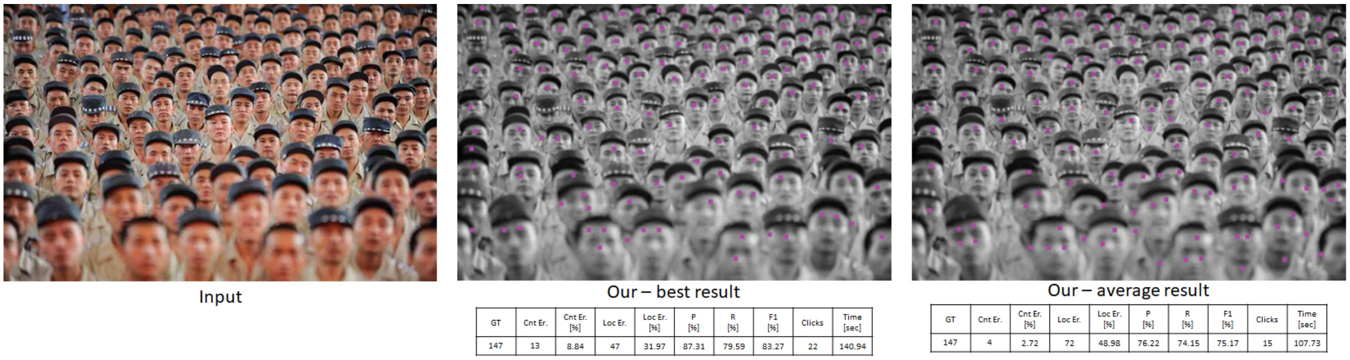
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
260	14	5.38	54	20.77	87.59	92.31	89.89	18	336.14

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
260	23	8.85	67	25.77	84.1	91.54	87.66	26	168.8

Figure 30: Sheep.



Input

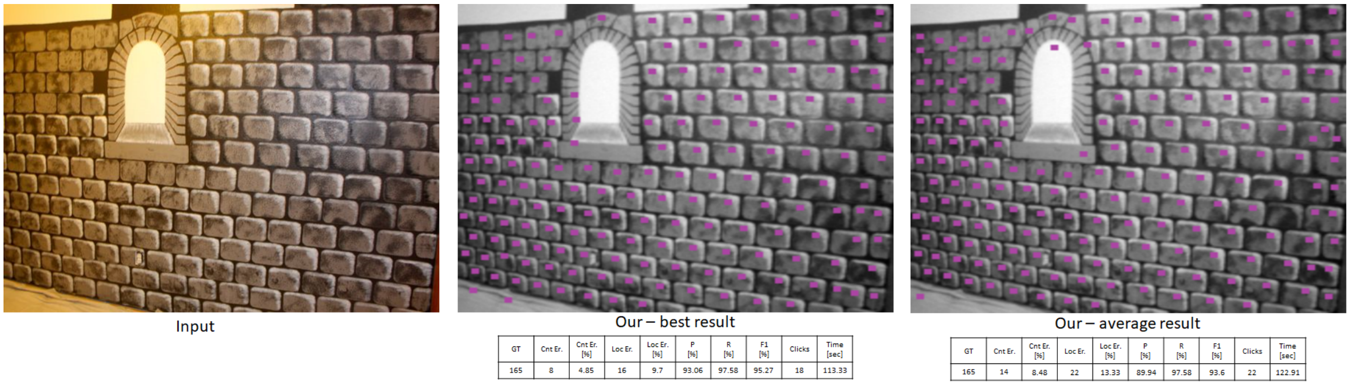
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
147	13	8.84	47	31.97	87.31	79.59	83.27	22	140.94

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
147	4	2.72	72	48.98	76.22	74.15	75.17	15	107.73

Figure 31: Soldiers. Shanghaitech dataset [6].



Input

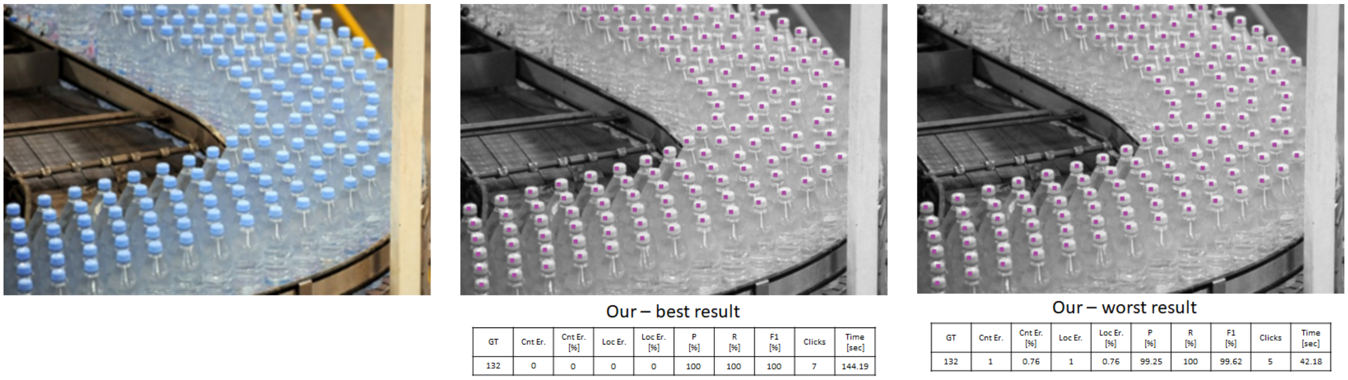
Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
165	8	4.85	16	9.7	93.06	97.58	95.27	18	113.33

Our – average result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
165	14	8.48	22	13.33	89.94	97.58	93.6	22	122.91

Figure 32: Wall.



Our – best result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
132	0	0	0	0	100	100	100	7	144.19

Our – worst result

GT	Cnt Er.	Cnt Er. [%]	Loc Er.	Loc Er. [%]	P [%]	R [%]	F1 [%]	Clicks	Time [sec]
132	1	0.76	1	0.76	99.25	100	99.62	5	42.18

Figure 33: Water.

References

[1] C. Arteta, V. Lempitsky, J. A. Noble, and A. Zisserman. Interactive object counting. In *European Conference on Computer Vision*, 2014. 9877, 9878

[2] Elena Bernardis and Stella Yu. Pop out many small structures from a very large microscopic image. *Medical image analysis*, 15:690–707, 07 2011. 9878, 9888

[3] Inbar Huberman and Raanan Fattal. Detecting repeating objects using patch correlation analysis. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2903–2911, 2016. 9877, 9878

[4] Antti Lehmussola, Pekka Ruusuvuori, Jyrki Selinummi, Heikki Huttunen, and Olli Yli-Harja. Computational framework for simulating fluorescence microscope images with cell populations. *IEEE Trans. Med. Imaging*, 26(7):1010–1016, 2007. 9878, 9882

[5] Z. Ma, Lei Yu, and A. B. Chan. Small instance detection by integer programming on object density maps. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3689–3697, June 2015. 9878, 9880, 9881, 9885

[6] Yingying Zhang, Desen Zhou, Siqin Chen, Shenghua Gao, and Yi Ma. Single-image crowd counting via multi-column convolutional neural network. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 589–597, 2016. 9878, 9888