

A. Appendix

A.1. Dataset

		Total	PMC	Spg Ad	Spg	Pl-Z	P	rsptd	esptd	SC
0	val	1,050	72	47	93	167	220	134	192	125
	train	4,150	283	184	366	659	877	531	759	491
1	val	1,047	72	47	93	167	220	134	190	124
	train	4,153	283	184	366	659	877	531	761	492
2	val	1,041	71	46	92	164	220	134	190	124
	train	4,159	284	185	367	662	877	531	761	492
3	val	1,036	71	46	91	164	220	132	190	122
	train	4,164	284	185	368	662	877	533	761	494
4	val	1,026	69	45	90	164	217	131	189	121
	train	4,174	286	186	369	662	880	534	762	495

Table A.1. 5-fold cross-validation split for training and validation dataset.

A.2. Self-Supervised Augmentation Settings

A.2.1 SimCLR

- colour jitter
 - probability = 0.8
 - strength = 1.0
 - brightness = 0.8
 - contrast = 0.8
 - saturation = 0.8
 - hue = 0.2
- random resized crop (scale $\in (0.08, 1.0)$)
- random gray scale (probability = 0.2)
- gaussian blur
 - probability = 0.5
 - $\sigma \in (0.1, 0.2)$
- horizontal flip (probability = 0.5)

Scan Setting	Backbone					
	ResNet18	ResNet50	ResNet101	ViT-T	ViT-S	ViT-B
20× SZP	8192	4096	3072	1024	512	256
40× EFI	4096	1024	768	1024	512	256

Table A.2. Batch sizes for SimCLR Training.

A.2.2 DINO

- 2 global views
 - size = image size
 - random resized crop (scale $\in (0.5, 1.0)$)
- 6 local views

- size = $\frac{\text{image size}}{2}$
- local random resized crop (scale $\in (0.1, 0.5)$)

- colour jitter
 - probability = 0.8
 - strength = 0.5
 - brightness = 0.8
 - contrast = 0.8
 - saturation = 0.4
 - hue = 0.2
- random gray scale (probability = 0.2)
- gaussian blur
 - probability (global view 1, global view 2, local views) = (1.0, 0.1, 0.5)
 - $\sigma \in (0.1, 0.2)$
- solarisation (probability = 0.2)
- horizontal flip (probability = 0.5)

Scan Setting	Backbone					
	ResNet18	ResNet50	ResNet101	ViT-T	ViT-S	ViT-B
20× SZP	8192	1536	1280	512	256	128
40× EFI	3072	768	512	512	256	128

Table A.3. Batch sizes for DINO Training.

A.2.3 MAE

- patch masking ratio = 0.75
- random resize crop (scale $\in (0.5, 1.0)$)
- random horizontal flip (probability = 0.5)

Scan Setting	Backbone		
	ViT-T	ViT-S	ViT-B
20× SZP	3072	2048	1536
40× EFI	3072	2048	128

Table A.4. Batch sizes for MAE Training.

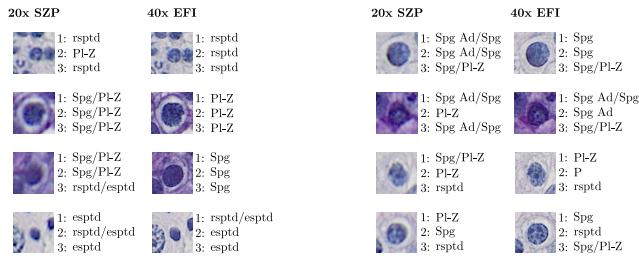


Figure A.1. Examples for annotation of some cell images in 20 \times SZP and 40 \times EFI quality by the three annotators. The numbers before the labels refer to the annotators (1-3).

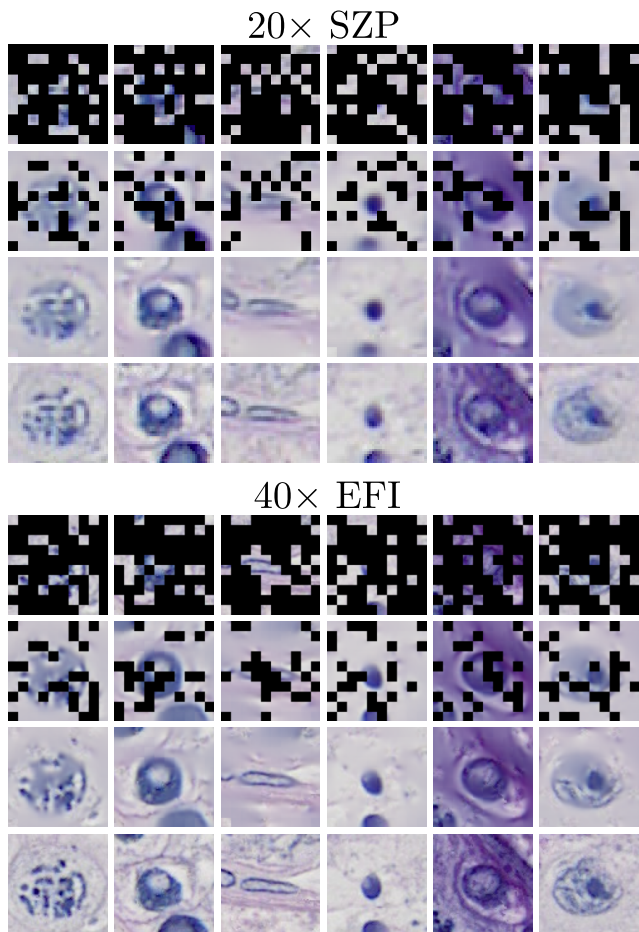


Figure A.2. ViT-B MAE reconstructions for 20 \times SZP (top) and 40 \times EFI (bottom) cell crops with background. From top to bottom the images display: 1. the input with masked areas in black, 2. the reconstructions with unmasked patches in black, 3. the combination of reconstructions and input patches, 4. the original image.

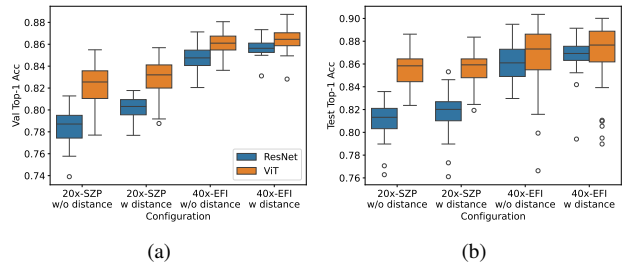


Figure A.3. Influence of distance information and scan settings on val (a) and test (b) Top-1 Acc on ResNet (18, 50, 101) and ViT (T,S,B) pretrained with any SSL approach. 300 models were trained only with background in 60 configurations. The legend of (a) also applies to (b).

Scan-Setting	Pretraining	Backbone	Background	Dist-Info	Data-Split	best fold Top-1 Acc (based on Val)	Mean Top-1 Acc	Mean Top-2 Acc	Mean Adjacent-1 Acc	Mean Adjacent-1 Dynamic Acc
20x-SZP	ImageNet	ResNet18	with	False	val test	0.825 0.833	0.820 ±0.004 0.827 ±0.014	0.940 ±0.003 0.955 ±0.009	0.921 ±0.003 0.953 ±0.005	0.932 ±0.005 0.973 ±0.004
20x-SZP	ImageNet	ResNet18	with	True	val test	0.836 0.846	0.826 ±0.008 0.832 ±0.015	0.945 ±0.006 0.963 ±0.006	0.926 ±0.006 0.957 ±0.009	0.934 ±0.006 0.974 ±0.006
20x-SZP	ImageNet	ResNet18	without	False	val test	0.814 0.808	0.802 ±0.009 0.800 ±0.010	0.931 ±0.004 0.939 ±0.005	0.907 ±0.005 0.930 ±0.006	0.916 ±0.007 0.948 ±0.006
20x-SZP	ImageNet	ResNet18	without	True	val test	0.826 0.844	0.822 ±0.004 0.824 ±0.014	0.943 ±0.006 0.957 ±0.007	0.924 ±0.005 0.952 ±0.003	0.934 ±0.002 0.972 ±0.006
20x-SZP	ImageNet	ResNet50	with	False	val test	0.828 0.836	0.821 ±0.008 0.830 ±0.008	0.942 ±0.004 0.958 ±0.007	0.921 ±0.005 0.954 ±0.006	0.930 ±0.006 0.970 ±0.005
20x-SZP	ImageNet	ResNet50	with	True	val test	0.831 0.856	0.826 ±0.003 0.839 ±0.012	0.949 ±0.007 0.969 ±0.008	0.929 ±0.006 0.957 ±0.005	0.939 ±0.009 0.976 ±0.007
20x-SZP	ImageNet	ResNet50	without	False	val test	0.821 0.783	0.806 ±0.012 0.811 ±0.017	0.929 ±0.006 0.945 ±0.009	0.911 ±0.011 0.937 ±0.010	0.922 ±0.011 0.955 ±0.011
20x-SZP	ImageNet	ResNet50	without	True	val test	0.827 0.843	0.822 ±0.005 0.835 ±0.011	0.946 ±0.005 0.965 ±0.006	0.927 ±0.007 0.952 ±0.007	0.934 ±0.009 0.967 ±0.008
20x-SZP	ImageNet	ResNet101	with	False	val test	0.823 0.831	0.819 ±0.004 0.811 ±0.016	0.942 ±0.003 0.960 ±0.008	0.924 ±0.007 0.952 ±0.010	0.932 ±0.009 0.964 ±0.010
20x-SZP	ImageNet	ResNet101	with	True	val test	0.839 0.851	0.830 ±0.010 0.835 ±0.016	0.946 ±0.010 0.969 ±0.005	0.930 ±0.005 0.957 ±0.009	0.938 ±0.008 0.969 ±0.008
20x-SZP	ImageNet	ResNet101	without	False	val test	0.816 0.820	0.806 ±0.009 0.802 ±0.010	0.931 ±0.008 0.943 ±0.007	0.913 ±0.007 0.926 ±0.005	0.922 ±0.006 0.943 ±0.007
20x-SZP	ImageNet	ResNet101	without	True	val test	0.828 0.824	0.822 ±0.009 0.829 ±0.006	0.946 ±0.009 0.962 ±0.002	0.924 ±0.012 0.949 ±0.004	0.929 ±0.014 0.965 ±0.008
40x-EFI	ImageNet	ResNet18	with	False	val test	0.861 0.826	0.857 ±0.004 0.825 ±0.013	0.962 ±0.008 0.951 ±0.010	0.944 ±0.003 0.951 ±0.011	0.947 ±0.005 0.963 ±0.011
40x-EFI	ImageNet	ResNet18	with	True	val test	0.867 0.852	0.864 ±0.002 0.848 ±0.014	0.963 ±0.004 0.963 ±0.005	0.950 ±0.003 0.959 ±0.004	0.957 ±0.004 0.973 ±0.003
40x-EFI	ImageNet	ResNet18	without	False	val test	0.854 0.737	0.846 ±0.008 0.751 ±0.030	0.954 ±0.005 0.896 ±0.026	0.933 ±0.004 0.904 ±0.023	0.937 ±0.006 0.923 ±0.023
40x-EFI	ImageNet	ResNet18	without	True	val test	0.856 0.760	0.853 ±0.003 0.779 ±0.019	0.960 ±0.004 0.923 ±0.013	0.941 ±0.003 0.925 ±0.017	0.947 ±0.002 0.945 ±0.010
40x-EFI	ImageNet	ResNet50	with	False	val test	0.864 0.870	0.857 ±0.007 0.845 ±0.028	0.959 ±0.007 0.958 ±0.012	0.947 ±0.003 0.957 ±0.011	0.951 ±0.004 0.959 ±0.012
40x-EFI	ImageNet	ResNet50	with	True	val test	0.872 0.862	0.864 ±0.005 0.853 ±0.012	0.963 ±0.005 0.970 ±0.007	0.948 ±0.005 0.959 ±0.006	0.954 ±0.007 0.969 ±0.005
40x-EFI	ImageNet	ResNet50	without	False	val test	0.849 0.778	0.842 ±0.005 0.776 ±0.005	0.953 ±0.009 0.930 ±0.008	0.932 ±0.005 0.921 ±0.005	0.938 ±0.008 0.930 ±0.008
40x-EFI	ImageNet	ResNet50	without	True	val test	0.865 0.810	0.856 ±0.008 0.811 ±0.006	0.962 ±0.003 0.941 ±0.011	0.943 ±0.004 0.940 ±0.007	0.951 ±0.005 0.952 ±0.012
40x-EFI	ImageNet	ResNet101	with	False	val test	0.863 0.839	0.859 ±0.003 0.840 ±0.003	0.958 ±0.005 0.959 ±0.009	0.945 ±0.005 0.955 ±0.006	0.950 ±0.005 0.961 ±0.005
40x-EFI	ImageNet	ResNet101	with	True	val test	0.873 0.850	0.866 ±0.009 0.839 ±0.012	0.962 ±0.008 0.960 ±0.005	0.950 ±0.007 0.954 ±0.005	0.957 ±0.006 0.960 ±0.009
40x-EFI	ImageNet	ResNet101	without	False	val test	0.849 0.801	0.839 ±0.008 0.786 ±0.011	0.953 ±0.002 0.927 ±0.011	0.930 ±0.006 0.924 ±0.012	0.939 ±0.006 0.944 ±0.007
40x-EFI	ImageNet	ResNet101	without	True	val test	0.860 0.798	0.852 ±0.008 0.817 ±0.017	0.960 ±0.003 0.948 ±0.007	0.941 ±0.007 0.947 ±0.008	0.948 ±0.008 0.958 ±0.012

Table A.6. Aggregated baseline results pretrained on ImageNet evaluating different scan qualities (20× SZP and 40× EFI), usage of cell background and distance information on validation and test dataset. Results in **bold** indicate the best validation results, and additionally, underlined values are the best test results within sections separated by double lines.

