Supplemental Materials

Molecular Biology of the Cell

Nagao et al.

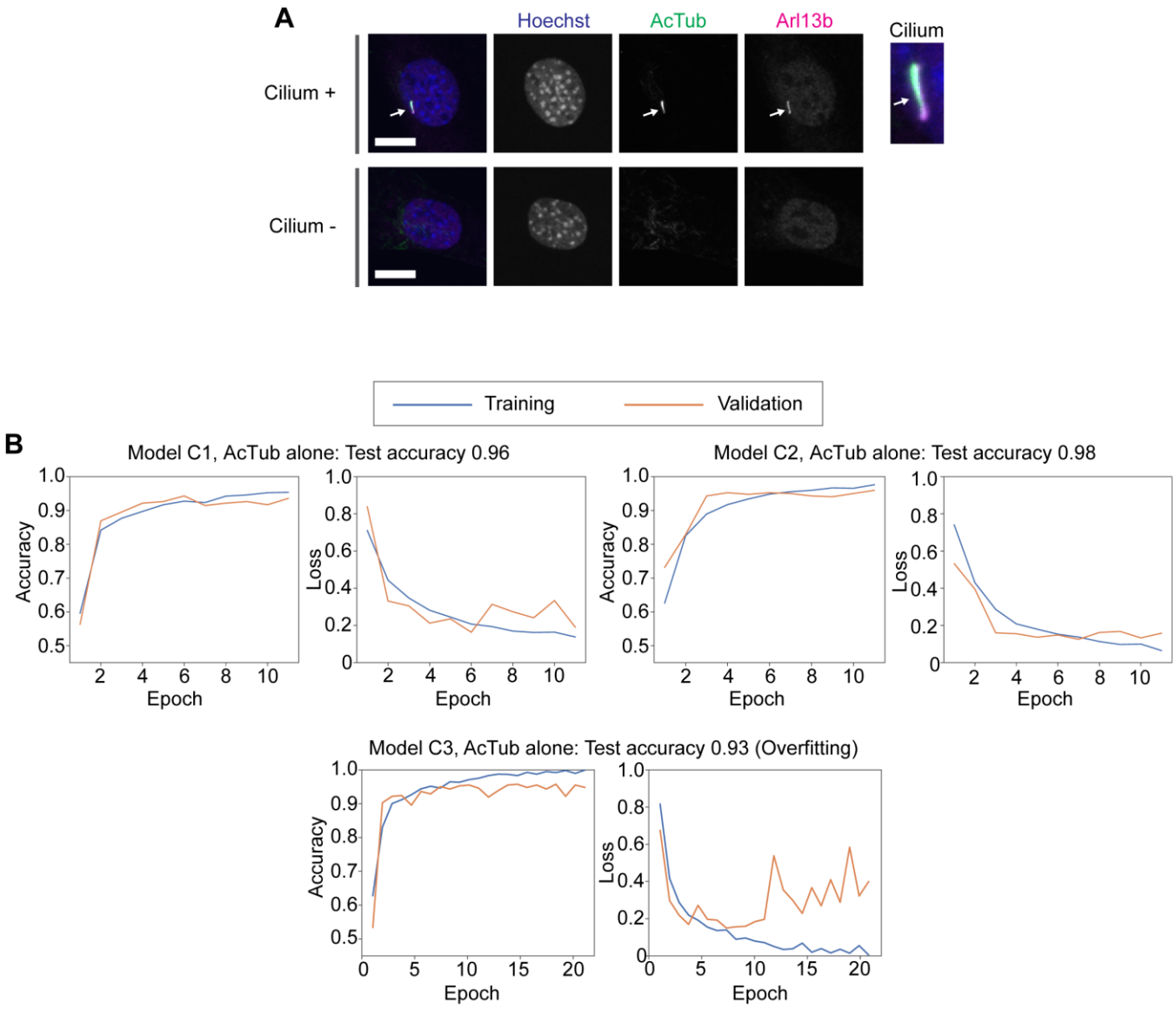


FIGURE S1: Classification of ciliated and non-ciliated cells by CNN. **(A)** Representative images of ciliated (Cilium +) and non-ciliated (Cilium -) NIH3T3 cells. The cilium is pointed with an arrow and magnified on the right-most (Cilium). **(B)** Learning curves of two representative models. The accuracies of the test data are shown above each graph.

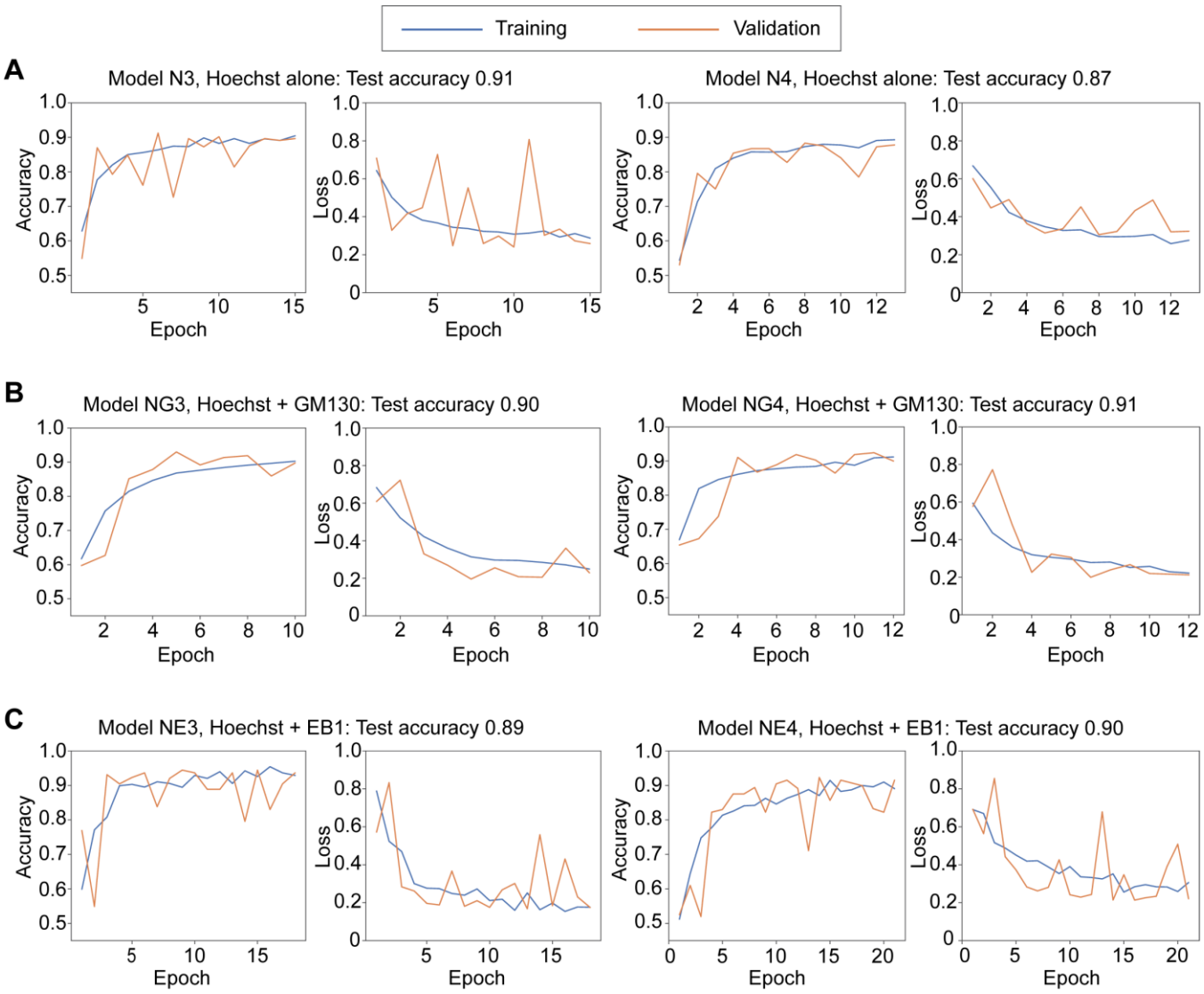


FIGURE S2: Additional representative learning curves of CNN models. Learning curves of two more representative models in each condition besides those in Fig 2 are shown.

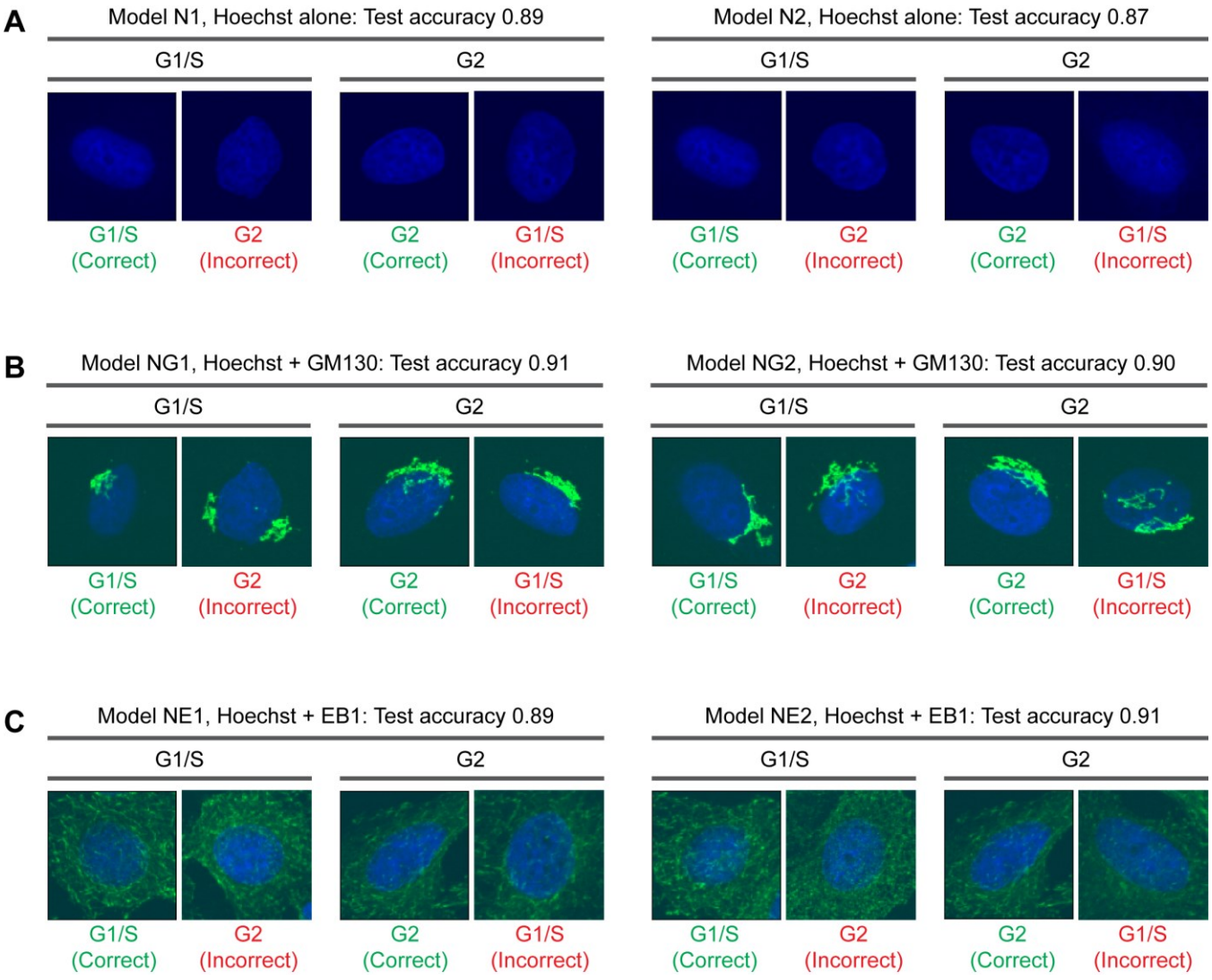


FIGURE S3: Images classified with CNN models. Representative images with class labels generated by CNN models. Both correctly and incorrectly classified images are shown for each category. Annotation labels (G1/S or G2) are shown on the top of images. Labels on the bottom of each image indicate class labels generated by CNN models (green for correct and red for incorrect). The models correspond to those in Fig 2.

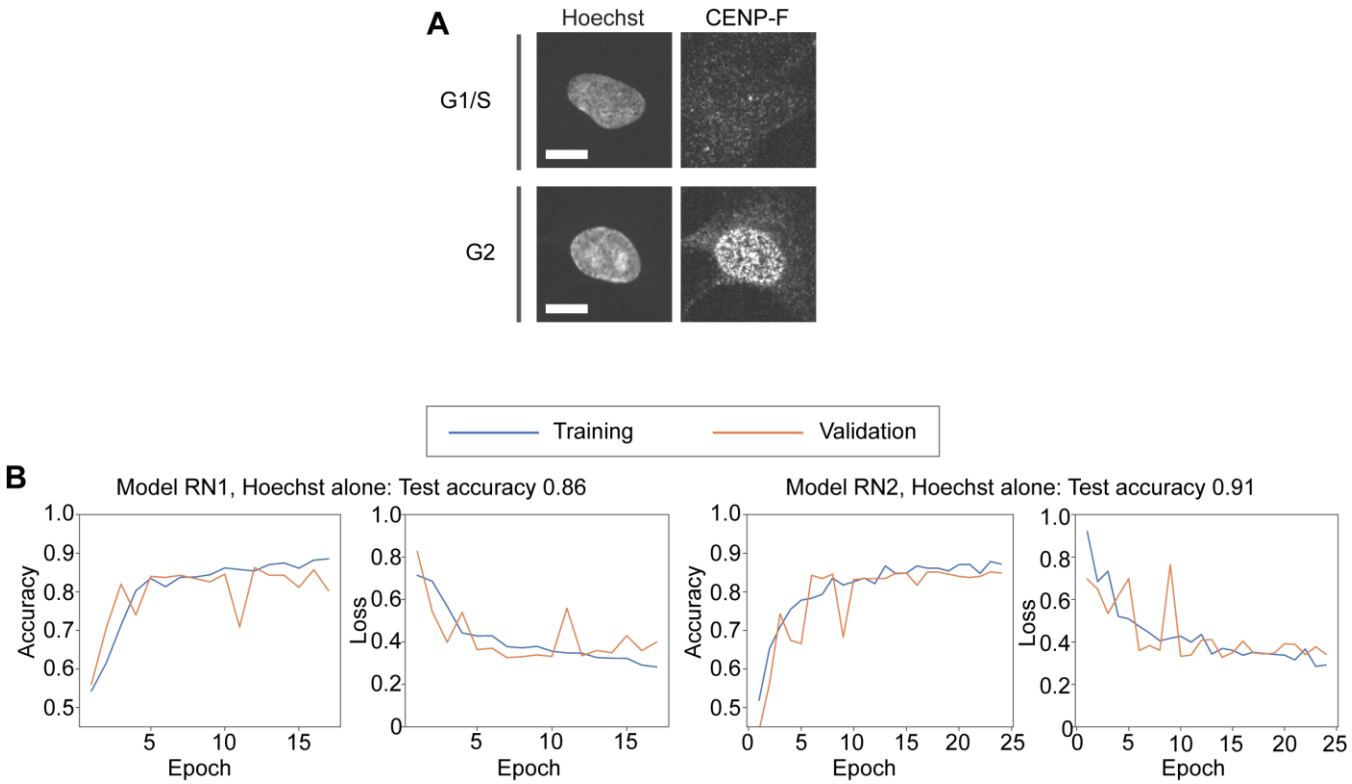


FIGURE S4: CNN-based classification of G1/S and G2 phases using Hoechst staining of RPE1 cells. **(A)** Representative images of RPE1 cells stained with Hoechst and an antibody to CENP-F. **(B)** Learning curves of two representative models.

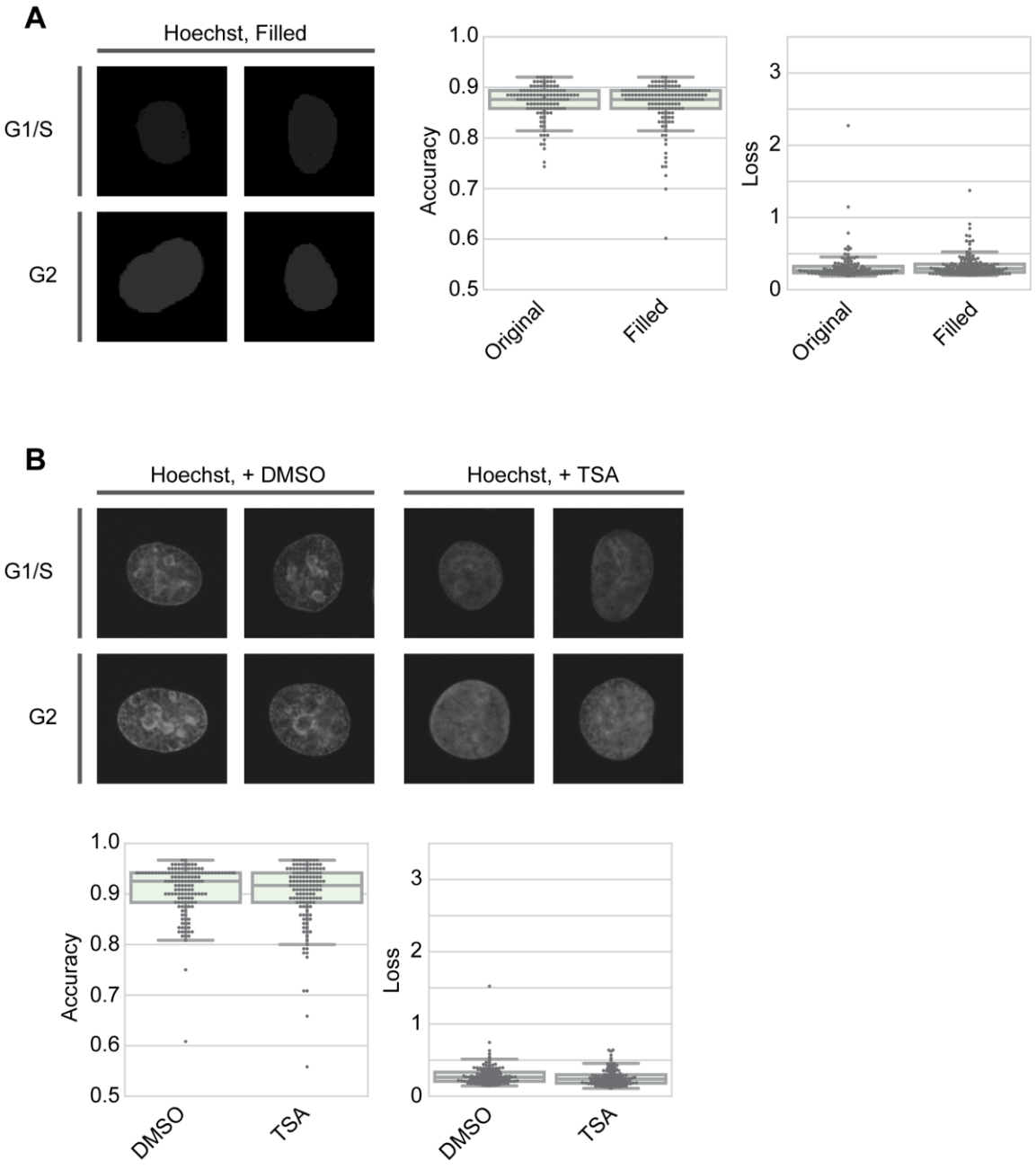
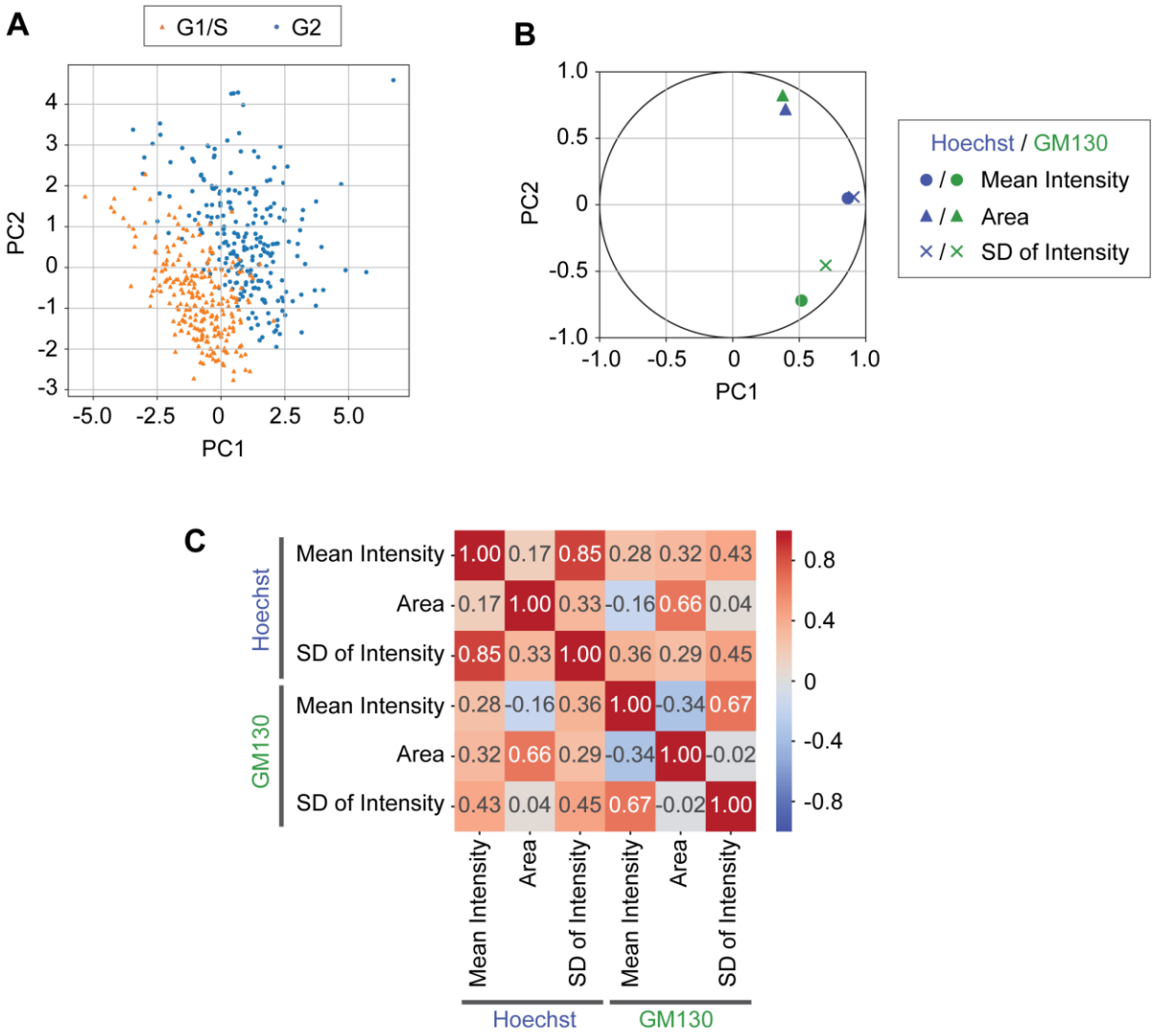


FIGURE S5: Effect of removal of nuclear texture on the classification. **(A)** Effect of uniform filling of nuclei on the classification. The images of HeLa cells stained with Hoechst were binarized to fill the nuclei with their mean intensity values and the background with 0. Representative images (left) and the results of Bayesian optimization (center, right) are shown. No significant difference was detected by Mann-Whitney U test. **(B)**

Effect of TSA treatment on the classification. HeLa cells were treated with DMSO (control) or TSA, followed by construction of CNN models. Representative images (top) and the results of Bayesian optimization (bottom) are shown. No significant difference was detected by Mann-Whitney U test.



D

		All Parameters					Selected Parameters		
		Coefficient	SE	Z	p value	VIF	Coefficient		
Hoechst	Mean Intensity	4.36	0.68	6.38	0.00	4.65	4.51	3.11	2.65
	Area	2.11	0.38	5.48	0.00	2.22	2.16	2.45	
	SD of Intensity	0.46	0.59	0.79	0.43	4.96			
GM130	Mean Intensity	-1.40	0.40	-3.47	0.00	2.44	-1.55		
	Area	1.32	0.45	2.90	0.00	2.47	1.15		2.79
	SD of Intensity	-0.40	0.32	-1.25	0.21	2.12			
	Train AUC	0.99					0.99	0.97	0.97
	Test AUC	0.98					0.98	0.96	0.97

FIGURE S6: Multivariate analyses of parameters from Hoechst-GM130 images to support the conclusions from the data in Fig 4. SD, standard deviation. **(A–B)** PCA of the six parameters. Plots of principal component scores

(A) and loadings **(B)** for the first two principal components (PC1 and PC2) are shown. **(C)** Correlation coefficient matrix of the six parameters. **(D)** Summary of logistic regression analyses. Coefficient, regression coefficient; SE, standard error of the coefficients; Z and p value, parameters of Wald test; VIF, variance inflation factor; AUC, area under the curve.

Supplementary File 1: Information for the Python package versions, the datasets, and CNN models, summarized in an Excel file.