Unbiased identification of cell identity in dense mixed neural cultures.

Sarah De Beuckeleer[1], Tim Van De Looverbosch[1], Johanna Van Den Daele[1], Peter Ponsaerts[2], and Winnok H. De Vos[1,3,4]

[1] Laboratory of Cell Biology and Histology, University of Antwerp

[2] Laboratory of Experimental Hematology, Vaccine and Infectious Disease Institute (Vaxinfectio), University of Antwerp

[3] Antwerp Centre for Advanced Microscopy, University of Antwerp

[4] µNeuro Research Centre of Excellence, University of Antwerp

e-mail:

sarah.debeuckeleer@uantwerpen.be

tim.vandelooverbosch@uantwerpen.be

johanna.vandendaele@uantwerpen.be

peter.ponsaerts@uantwerpen.be

winnok.devos@uantwerpen.be

Contact information *Corresponding author:

University of Antwerp, Laboratory of Cell Biology and Histology

Universiteitsplein 1, 2610 Antwerp, Belgium

e-mail: winnok.devos@uantwerpen.be

Keywords: morphological phenotyping, multiplexed imaging, iPSC culture validation, neural differentiation, computational biology

Abstract

Induced pluripotent stem cell (iPSC) technology is revolutionizing cell biology. However, the variability between individual iPSC lines and the lack of efficient technology to comprehensively characterize iPSC-derived cell types hinder its adoption in routine screening settings. To facilitate the validation of iPSC-derived cell culture composition, we have implemented an imaging assay based on cell painting and convolutional neural networks to recognize neural cell types in dense and mixed cultures with high fidelity. We have benchmarked our approach using pure and mixed cultures of neuroblastoma and astrocytoma cell lines and attained a classification accuracy above 96%. Through iterative data erosion we found that inputs containing the nuclear region of interest and its close environment, allow achieving equally high classification accuracy as inputs containing the whole cell for semi-confluent cultures and preserved its accuracy even in very dense cultures. We then applied the nucleocentric cell profiling approach to evaluate the differentiation status of iPSC-derived neural cultures, by determining the ratio of postmitotic neurons and neural progenitors. We found that the cell-based prediction significantly outperformed an approach in which the culture time point was used as classification criterion (96% *vs.* 86%, resp.). Thus, nucleocentric morphological single cell profiling provides a means to quantify cell composition in complex mixed neural cultures and holds promise for use in quality control of iPSC-derived neural cell culture models.

Introduction

Modelling the complexity of the human brain and its (dys)function has proven to be notoriously challenging. This is due to its intricate wiring, the cellular heterogeneity and species-specific differences[1]. To increase the translational value of neuroscientific research, there is a need for physiologically relevant human models. With the advent of human induced pluripotent stem cell (iPSC) technology, it has become possible to generate a wealth of brain-resident cell types including neurons[2], astrocytes[3], microglia[4], oligodendrocytes[5] and endothelial cells[6], allowing the study of complex polygenic pathologies that cannot be modelled in animals, opening avenues to precision pharmacology[7,8]. Furthermore, the ability of iPSC to develop into organoids renders them attractive for studying multi-cellular interactions in a 3D context that is closer to the *in vivo* situation[9]. However, genetic drift and batch-to-batch heterogeneity cause variability in reprogramming and differentiation efficiency[10]. This hinders the use of iPSC-derived cell systems in systematic drug screening or cell therapy pipelines. Quality control in such settings demands a robust and cost-effective approach. Current culture validation methods include sequencing, flow cytometry and immunocytochemistry [10,11]. These methods are often low in throughput, costly and/or destructive. Thence, we set out to develop a method for evaluating the composition of such cultures based on high-content imaging, which is fast and affordable. The primary criteria were to facilitate cell type identification with cellular resolution while preserving spatial information and ensuring compatibility with subsequent immunocytochemistry or molecular assays for further biological inquiries. To this end, we employed the Cell Painting (CP)[12,13] method, which is based on labelling cells with simple organic dyes and analysing the resulting phenotypes. CP has proven to be a powerful and generic method for predicting the mode-of-action of pharmacological or genetic perturbations, and this sheerly using a cell morphology readout[14–16]. Thus far, CP has primarily been utilized to predict conditions associated with pharmacological treatments or genetic modifications

using images as input. An exemplary achievement in this context has been realized by the JUMP consortium, which has published an open-source CP image dataset comprising over 13,000 distinct conditions[17]. Recognizing the predictive potential of morphological information, we explored its utility at the single-cell level, shifting focus from the image level to predict individual cell types. By implementing deep learning for cell classification and tuning the input space, we established an approach that allows recognizing cell types with high accuracy, even in very dense cultures. We illustrate its potential by using it to distinguish iPSC-derived cell types in co-culture and to assess the maturity of differentiating neurons.

Results

Neural cell lines have a unique morphotextural fingerprint.

Several groups have demonstrated that morphological cell profiling can be used to discriminate pharmacogenomic perturbations based on phenotypic similarity[14–16]. We asked whether a similar approach could be leveraged to unequivocally distinguish individual cell types as well. To this end, we implemented a version of CP based on 4-channel confocal imaging (**Fig. 1A**) and applied it to monocultures of two neural cell lines from a different lineage, namely astrocyte-derived 1321N1 astrocytoma cells and neural crest-derived SH-SY5Y neuroblastoma cells. First, we explored whether traditional morphotextural feature extraction from cells and their nuclei provided sufficient distinctive power. Representation of the resulting standardized feature set in UMAP space revealed a clear separation of both cell types without replicate bias (**Fig. 1B**). Clustering of instances was less pronounced after principal component analysis (**Suppl. Fig. 4 A**), but UMAP better preserves local and global data structure [18]. When mapping the individual features to the UMAP space, we found that both texture (*e.g.,* energy, homogeneity) and shape (*e.g.,* nuclear area, cellular area) metrics contributed to the separation (**Fig. 1C**). The contribution of intensity-related features (*e.g.,* mean, max, min, std intensity) to cell type separation was less pronounced as they were more correlated with the biological replicate. Thus, we concluded that cell types can be separated across replicates based on a morphotextural fingerprint.

CNN outperforms random forest in cell type classification.

To evaluate whether the morphotextural fingerprint could be used to predict cell type, we performed Random Forest (RF) classification, using different seeds for splitting up training and validation sets. This resulted in a rather poor prediction accuracy of 71,0±1,0%, mainly caused by the significant (46%) misclassification of 1321N1 cells (**Fig. 1D**). The imbalance in

105 recall and precision was surprising given the clear separation of cell populations in UMAP
106 space using the same feature matrix. When inspecting the main contributions to the RF
107 classifier, we found very similar features as highlighted in UMAP space to add to the
108 discrimination (**Fig. 1C**). Where the top 2 features (nuclear Cy3 energy and cellular area)
109 showed the expected gradient along the UMAP1 direction, the third parameter (cellular DAPI
110 contrast) had no contribution to UMAP separation. Reducing noise by manually removing
111 redundant and correlating features could not ameliorate RF performance, which may be due to
112 the documented bias in feature selection for node splitting in high-dimensional data[19]. This
113 result drove us to improve the classification accuracy by means of a ResNet[20] convolutional
114 neuronal network (CNN). Here, we no longer relied on the extraction of features from
115 segmented cell objects, but instead used image crops centred around individual cells (bounding
116 box around cell centroid) and blanked for their surroundings as input (**Fig. 2A**). Using this
117 approach, we found a significantly higher prediction accuracy of 96,0±1,8%, with a much more
118 balanced recall and precision (**Fig. 2B**). Even with only 100 training instances per class, CNN
119 outperformed RF, but optimal performance was attained with 5000 training instances (**Fig.
120 2C**). As expected, a model trained on all three biological replicates performed slightly worse
121 (lower accuracy and higher variability) at predicting a test set of a given replicate than a model
122 trained on that same replicate (**Fig. 2D**). However, models trained on two different replicates
123 outperformed those that were only trained on one when used for predicting instances of a third
124 (unseen) replicate. This emphasizes the need for including sufficient variation in the training
125 set (**Fig. 2E**). Although more performant, CNN classification does not allow direct retrieval of
126 discriminative features, which complicates model interpretation. To gain a visual
127 understanding of image information contributing most to the classification we therefore
128 resorted to Gradient-weighted Class Activation Mapping (Grad-CAM)[21]. This revealed that the
129 cell borders (edge information), nuclear and nucleolar signals were the most distinctive features

130 (**Fig. 2F**). When scrutinizing CNN misclassifications, we found that these are mainly caused
131 by faulty cell detection (*e.g.,* oversegmentation of cell ramifications), unhealthy/dead cells,
132 debris or visibly aberrant cell shapes (**Suppl. Fig. 4B**) - mistakes not necessarily attributed to
133 the CNN. To shed light on the contribution of individual markers to the classification, we
134 eroded the input to single channel images. For all cases, the prediction performance was below
135 or equal to 85,0% suggesting that no single channel contains all relevant information (**Fig. 2**
136 **G**). Combinations of two or three channels could not match the prediction accuracy of the full
137 four-channel image either. Thus, we concluded that all channels contribute at least in part to a
138 successful classification.
139
140 *Nucleocentric predictions remain accurate regardless of culture density.*
141 Morphological profiling relies on cell identification for adequate class prediction. This may
142 become difficult in dense cultures such as iPSC-cultures, clustered cells and tissue (mimics).
143 Having established a method to distinguish cell types with high accuracy, we next asked how
144 robust the classification was to increasing cell density. To this end, we acquired images of
145 1321N1 and SH-SY5Y monocultures, grown at densities ranging from 0 to 100% confluency
146 (**Fig. 3A**). Based on the nuclear count, we binned individual fields into 6 density classes (0-
147 20%, 20-40%, 40-60%, 60-80%, 80-95%, 95-100%) and trained a CNN with equal sampling
148 of cell numbers per density class to avoid bias. No decrease in accuracy was observed until
149 culture density reached 80% confluency. Only for very high densities (95-100%), we found a
150 significant decrease in the prediction accuracy (92,0±1,7%) (**Fig. 3B**). We reasoned that under
151 these conditions cell shape would be predominantly determined by neighbouring cells and cell
152 segmentation performance would decrease (**Suppl. Fig. 1**). Nuclei are less malleable and even
153 in dense cultures remain largely separated, allowing their robust segmentation and avoiding
154 CNN misclassifications resulting from segmentation errors. Hence, we asked whether using

155 the nuclear ROI as input would improve classification performance at high densities despite
156 containing less cellular information. However, regardless of a relatively high average
157 prediction accuracy of about 92,4±1,0% (**Fig. 3C**), the performance was lower than cell ROI
158 across the density range. Thence, we tested an intermediate condition in which nuclear ROI
159 with their direct surroundings (a square bounding box of 60x60 microns surrounding the
160 nuclear centroid coordinates) were considered as input (**Suppl. Fig. 4C**). This resulted in a
161 classification accuracy of 95,0±0,4% (**Fig. 3C**) comparable to that of models trained on whole-
162 cell inputs for low to moderate cell densities, but also maintained prediction performance at
163 high cell densities (**Fig. 3B**). To understand why this was not the case for nuclear crops, we
164 inspected Grad-CAM output for these predictions, and found that the latter diverted the
165 attention of the CNN to the background (**Fig. 3D**). We interpret this result as the CNN using
166 the background as a proxy for nuclear size. As nuclear area dropped with culture density, the
167 dynamic range decreased, which could explain the increased error rate of the CNN for high
168 densities unrelated to segmentation errors (**Suppl. Fig. 4B**). Thus, we conclude that using the
169 nucleocentric region as input for the CNN is a valuable strategy for accurate cell phenotype
170 identification in dense cultures.
171
172 *Cell prediction remains accurate in mixed cell culture conditions.*
173 Although a very high classification accuracy was obtained with nucleocentric CNN
174 predictions, both training and testing were performed with input images drawn from
175 monocultures. As our goal was to allow cell type prediction in complex, heterogeneous
176 cultures, we next switched to a more faithful situation in which we co-cultured both cell types.
177 Ground-truth for these predictions was generated by either performing post-hoc
178 immunofluorescence (IF) with cell-type specific antibodies (CD44 for 1321N1 and TUBB3 for
179 SH-SY5Y cells), or by differential replication labelling (EdU or BrdU for either of the cell

180 types) after dye quenching[22] (**Fig. 4A**). The latter proved significantly more successful for
181 binary ground truth classification (through intensity thresholding) than the former (**Fig. 4B**).
182 When training a CNN to recognize the cell types using these post-hoc markers as ground truth,
183 we found that the prediction accuracy on the left-out dataset of the co-culture (Co2Co) was
184 almost as high as when a model was trained and tested on monocultures (Mono2Mono) (**Fig.**
185 **4C**).
186 We then tested whether it was possible to use a classifier trained on monocultures for cell type
187 prediction in co-culture (Mono2Co). This resulted in a prediction accuracy of 85,0±4,6%. We
188 contributed this drop in accuracy to the effect on cell phenotype exerted by the presence of
189 other cell types in mixed cultures which is not captured in monocultures. As the images from
190 monocultures and co-cultures were obtained from different plates, we suspected inter-replicate
191 variability in culture and staining procedures to contribute in part to this lesser performance.
192 Therefore, we tested whether we could improve the performance of the CNN by including
193 monocultures from the same plate as the co-cultures. This finetuning indeed improved the
194 average performance to 89,0±0,7%, and more importantly, it significantly reduced the
195 variability (coefficient of variation) of the predictions (**Fig. 4D**). Thus, while not yet reaching
196 the same accuracy, it proves that it is possible to establish a model that recognizes cell types in
197 co-cultures that is solely trained on monocultures.
198
199 *Nucleocentric phenotyping can be applied to stage iPSC-derived neuronal cultures.*
200 iPSC-derived neuronal cell cultures suffer from significant inter- and intra-batch variability
201 and could benefit from an efficient quality control[23,24]. Thence, we applied our nucleocentric
202 phenotyping pipeline to stage the maturity of a neuronal cell culture based on its cell type
203 composition. Using a guided protocol[2], two differentiation stages were simulated: a primed
204 stage, where most cells are cycling neural progenitors (NPCs), and a differentiated stage where

205 most cells are post-mitotic neurons (**Fig. 5A**). The two cell types were discriminated by post-
206 hoc IF labelling for the cell cycle marker Ki67 (NPC) and the microtubule marker ß-III-tubulin
207 (TUBB3, neurons). Not all cells in the CP image could be assigned with a ground truth label
208 due to cell detachment upon re-staining or sheer absence of the tested marker. Since no single
209 monoculture consists of 100% of either cell type, we applied gates to retain only those cells
210 that show unequivocal staining for either one of both markers. Based on these gates, ROIs were
211 either classified as neuron (Ki67-/TUBB3+), NPC (Ki67+/TUBB3-) or undefined (outside of
212 gating boundaries). We assume the latter category to represent transitioning cells in
213 intermediate stages of neural development, un- or dedifferentiated iPSCs. (**Fig. 5B**). This
214 gating strategy resulted in a fractional abundance of neurons vs. NPC of 36,4 % in the primed
215 condition and 80,0% in the differentiated condition (**Fig. 5C**). UMAP dimensionality reduction
216 revealed a clear clustering of both phenotypes suggesting that the morphotextural profile
217 captured their distinct phenotypic fingerprint and should enable their separation (**Fig. 5D**).
218 In a first attempt to classify cells within the guided culture, we trained a CNN on individual
219 cell inputs using the culture condition as a weak label. This resulted in a prediction accuracy
220 of 87,0±0,9%, which is not unexpected given the cellular heterogeneity within each condition.
221 When we used the IF ground truth labels instead, we obtained a classification performance of
222 96,0±0,5 % (**Fig. 5E**). Applying this cell-based (as opposed to condition-based) CNN to the
223 two culture conditions resulted in predicted fractional abundance of neurons to NPC of 40,5%
224 in images of the primed condition and 74,2% in images of the differentiated condition –
225 aligning well with the manually defined ratios (**Fig. 5F**).
226 We then went on to evaluate the established cell-based classifier to a primed neuronal cell
227 culture undergoing gradual spontaneous differentiation after dual-SMAD inhibition[25]. We
228 examined cell cultures at 13, 30, 60 and 90 days *in vitro* (DIV) after the start of the
229 differentiation process and visually confirmed a gradual change in neural maturity, as

230 evidenced by the increasing dendritic outgrowth (**Fig. 5G**). In absence of post-hoc ground truth
231 labelling, we plotted the cells of the different DIVs onto the UMAP space of the guided
232 differentiation and quantified their relative abundance in regions encompassing either NPCs or
233 neurons (**Fig. 5H-I**). This confirmed a shift in the neuron-to-NPC fractional abundance with
234 an especially strong accumulation of neurons at DIV90. The cell-based CNN model trained on
235 the guided differentiation dataset recapitulated this shift in its predictions (**Fig. 5J**).
236
237 *Nucleocentric phenotyping can be applied to characterize mixed iPSC-derived neuronal*
238 *cultures*
239 Next to NPCs and neurons, iPSC-derived neuronal cultures are often studied in conjunction
240 with other relevant cell types that influence neuronal connectivity and homeostasis such as
241 astrocytes and microglia[26]. Therefore, we tested whether the cell-based approach could be
242 extended to these cell types as well. We generated iPSC-derived astrocytes, neurons, and
243 microglia from the same parental iPSC line (**Fig. 6A**) and trained a CNN to identify each cell
244 type. In monocultures, this led to a prediction accuracy of 96,8±1,0% (**Fig. 6B**) and in mixed
245 cultures of neurons and microglia (using NeuN, resp. Iba1 as ground-truth IF markers) (**Fig.**
246 **6C**) an accuracy of 96,9±0,5% (**Fig. 6D**). Based on these results, we conclude that
247 nucleocentric phenotyping can be used to gauge the composition of iPSC-derived cultures.
248

Discussion

iPSC-derived cell cultures can improve translatability of biomedical research but represent a challenge due to their higher variability and multicellular composition. In this research, we have presented a method to efficiently identify neural cell types in mixed cultures to aid in their validation and application in routine screening settings. We first benchmarked our approach using neural cell lines and found that a CNN-based approach outperforms shallow learning (RF) based on handcrafted feature extraction, even with a limited number of input images. This aligns well with recent data showing CNN-based feature representations outperformed shallow techniques for condition-based drug response prediction[27].

We found that all channels contribute to the overall prediction accuracy but that the whole cell is not required to obtain the highest prediction accuracy, especially in dense cultures. In line with earlier studies that highlight the biomarker potential of the nucleus for patient stratification[28], cell state identification[29] or drug response[14], we found that the nuclear ROI as such already carries highly distinctive information for cell type prediction. However, extension to its direct surroundings proved to be the best and most robust input across a range of cell densities. Since errors made by the CNN largely arose from segmentation errors, we hypothesize that such errors will be most pronounced in dense cultures. The nucleocentric approach, which is based on more robust nuclear segmentation, minimizes such mistakes. Furthermore, it opens possibilities for applying cell profiling to even more complex 3D cell systems such as tissue slices or organoids, where cell segmentation becomes extremely challenging.

One important conclusion from this work is that cell types can be identified in mixed cultures solely using the input from monocultures. This implies that cells retain their salient characteristics even in complex heterogeneous environments. While for now, predictions are still superior when based on training on mixed cultures, we found that the prediction accuracy

274 of monoculture trained models can be increased by employing replicate controls. This suggests
275 that it may become possible to apply or adapt existing models without the need for a cell-level
276 ground truth (as provided by post-hoc labelling). This technique could potentially be of use for
277 cultures or tissues where no antibody- or pre-labelling is possible (*e.g.,* no unique IF marker is
278 available, non-replicating cells). To increase the accuracy one could resort to intelligent data
279 augmentation[30] or transfer learning[31] strategies. While a ground-truth independent method
280 holds promise for bulk predictions (*e.g.,* for quality control purposes), the use of post-hoc
281 labelling allows building more refined models that can distinguish multiple cell types and/or
282 cell states at once. Especially using the cyclic staining that we have used[22], much richer
283 information content can be gained. With the method presented here, we combine the rich
284 multidimensional information from cyclic IF with morphological information and show that it
285 is possible to predict these IF outcomes. Multiplexed imaging has previously shown its utility
286 in gaining in depth information on culture composition and differentiation status in iPSC-
287 derived neural cultures (progenitor, radial glia, astrocytes (immature and mature) and neurons
288 (inhibitory and excitatory)[32]. Similarly, this could be expanded to cell states (apoptosis, DNA
289 damage)[15]. Initial model training requires both CP and multiplexed IF images. After the initial
290 training phase, the IF step can be omitted, and the CNN will predict the phenotype outcomes
291 of the IF with high accuracy on the CP images alone. This significantly reduces experimental
292 costs and time requirements.
293 Applying the method to iPSC-neuronal cultures revealed its potential to score their
294 differentiation state. Although guided protocols manage to speed up the differentiation process
295 and lead to enhanced culture purity, the neural differentiation process proves to be less than
296 binary, as evidenced by the mixed presence of Ki67+/TUBB3- and Ki67-/TUBB3+ cells. The
297 spontaneous differentiation protocol illustrated the unpredictable nature and late stage of the
298 differentiation process. Reports in literature highlight the difficulty of reproducible neural

differentiation and attribute this to culture conditions, cultivation time and variation in developmental signalling pathways in the source iPSC material[33,34]. Spontaneous neural differentiation has previously been shown to require approximately 80 days before mature neurons arise that can fire action potentials and show neural circuit formation. Although these differentiation processes display a stereotypical temporal sequence[25], the exact timing and duration might vary. Yet, we could reliably predict the gradual transition to a more differentiated state using our CNN model. Complementary, the same transition from NPC to neuronal phenotype could be observed in UMAP space. Although UMAP dimensionality reduction merely visualizes data structures and is less sensitive as compared to CNN classification, it aids in visualization and comprehension of the dataset. We realize that differentiating iPSC cultures are highly heterogenous and are composed of a landscape of transitioning progenitors in various stages of maturity that our binary models currently fail to capture. To predict more continuously the maturation of iPSC-derived neural progenitors to differentiated neurons, it could be of interest to consider deep learning-based regression approaches [35].

For now, we have tailored models to the individual datasets, but it is conceivable that a more generalist CNN could be established for multiple culture types. This would obviously demand a much larger dataset to encompass the variability encountered in such models (*e.g.,* various starting iPSC lines, various differentiation protocols). Publicly available datasets (*e.g.,* JUMP-Cell Painting Consortium) can aid in creating an input database containing a large variability (different iPSC lines, different neural differentiation protocols, …), which would ultimately lead to a more robust and generalist predictor. Our results showing the prediction accuracy of a guided differentiation model on spontaneously differentiating cultures indicate that the approach can be transferred to other differentiation protocols. Inclusion of more input images

323 and variability should thus enable developing a generalist model for other differentiation
324 protocols without the need for ground truth validation and further CNN training.
325 In conclusion, we present a novel application for unbiased morphological profiling by
326 extending its use to complex mixed neural cultures using sequential multispectral imaging and
327 convolutional network-informed cell phenotype classification. We show that the resulting
328 predictors are robust with respect to biological variation and cell culture density. This method
329 holds promise for use in quality control of iPSC cultures to allow their routine use in high-
330 throughput and high-content screening applications.
331

Methods

Cell culture. Cells were cultured at 37 °C and 5% CO2. 1321N1 and SH-SY5Y cell lines were maintained in DMEM-F12 + Glutamax (Gibco, 10565018) supplemented with 10% Fetal Bovine Serum (Gibco, 10500064). Cell seeding prior to imaging was done in 96-well black multiwell plates with #1.5 glass-like polymer coverslip bottom (Cellvis, P96-1.5P). Only the inner 60 wells were used, while the outer wells were filled with PBS-/- to avoid plate effects. Plates were coated with Matrigel (Corning, 734-1440) After seeding, the imaging plate was centrifuged at 100g for 3min.

iPSCs (Sigma Aldrich, iPSC0028 Epithelial-1) were cultured on Matrigel (Corning, 734-1440) in Essential 8 medium (Gibco, A1517001). Upon cell thawing, Rock inhibitor (Y-27632 dichloride, MedChem, HY-10583) was added in 10µM concentration. Subculturing of iPSCs was performed with ReLeSR (Stemcell Technologies, 05872).

Unguided differentiation[25] of iPS cells to neural progenitor cells (NPCs) was started by subculturing the iPSCs single cell using Tryple Express Enzyme (Life technologies, 12605010) at a density of 10e4 cells/cm² in mTesR1 medium (Stemcell Technologies, 85850) and Rock inhibitor. The following day, differentiation to NPCs was started by dual SMAD inhibition in neural maintenance medium (1:1 Neurobasal (Life technologies, 21103049):DMEM-F12+Glutamax (Gibco, 10565018), 0.5x Glutamax (Gibco, 35-050-061), 0.5% Mem Non Essential Amino Acids Solution (Gibco, 11140050), 0.5% Sodium Pyruvate (Gibco, 11360070), 50 µM 2-Mercaptoethanol (Gibco, 31350010), 0.025% Human Insulin Solution (Sigma Aldrich, I9278), 0.5X N2 (Gibco, 17502048), B27 (Gibco, 17504044), 50 U/ml Penicillin-Streptomycin (Gibco, 15140122)) supplemented with 1 µM LDN-193189 (Miltenyi, 130-106-540), SB431542 (Tocris, 1614). Daily medium changes were performed for 11 consecutive days. Following neural induction, neural rosettes were selected by STEMdiff Neural Rosette Selection Reagent (Stemcell technologies, 05832). Maintenance of neural

357 progenitor cells was performed in neural maintenance medium with 10 µM bFGF (Peprotech,
358 100-18C) added after subculturing. Single cell detachment of NPCs during maintenance was
359 performed using Tryple Express Enzyme. Cell seeding prior to imaging is done in 96-well
360 black multiwell plates µCLEAR (Greiner, 655090) coated with Poly-L-ornithine (Sigma-
361 Aldrich, P4957) and laminin (Sigma-Aldrich, L2020). Only the inner 60 wells were used, while
362 the outer wells were filled with PBS-/- to avoid plate effects. After seeding, the imaging plate
363 was centrifuged at 100g for 3min.
364 Guided iPSC differentiation to neurons was performed according to Bell et al. (2019)[2]. The
365 initial phase of neural induction consisted of 12 days neural induction medium (DMEM-
366 F12+Glutamax (Gibco, 10565018), 1x N2 (Gibco, 17502048), 1x B27 (Gibco,
367 17504044),1mg/ml BSA (Sigma-Aldrich, A7979), 0.5% Mem Non-Essential Amino Acids
368 Solution (Gibco, 11140050)). Of these 12 days, the first 7 were supplemented with cytokines
369 for dual-SMAD inhibition (1 µM LDN-193189 (Miltenyi, 130-106-540), SB431542 (Tocris,
370 1614)). Following neural induction, the NPCs were floated in uncoated MW6 culture plates in
371 NPC medium (DMEM-F12+Glutamax (Gibco, 10565018), 1x N2 (Gibco, 17502048), 1x B27
372 (Gibco, 17504044), 10 µM bFGF (Peprotech, 100-18C), 10 µM EGF (Peprotech, 100-47)).
373 After 4 days, NPC clusters of appropriate size were filtered using a cell strainer ((37 µm cell
374 strainer, Stemcell technologies, 27250) and plated on Matrigel-coated (Corning, 734-1440)
375 MW6 culture plates. NPCs can now be expanded in NPC medium. Guided differentiation into
376 forebrain neurons can be induced by switching to neuronal medium (BrainPhys (STEMCELL
377 Technologies, 05792), 1x N2 (Gibco, 17502048), 1x B27 (Gibco, 17504044), 10 µM BDNF
378 (PeproTech, AF-450-02), 10 µM GDNF (PeproTech, AF-450-02) for 15 days before fixation.
379 Differentiation of iPSC to microglia[4] was performed by the formation of embroid bodies (EBs)
380 with 10e3 iPSCs/well in a 96-well U-bottom plate (Corning, 351177) coated with Anti-
381 Adherence Rinsing Solution (Stemcell technologies, 07010) in mTeSR medium supplemented

382 with Rock inhibitor, 50 ng/mL BMP4 (Peprotech, 120-05), 50 ng/mL VEGF (Peprotech, 100-
383 20), 20 ng/mL SCF (Peprotech, 250-03). 75% medium is changed for 4 consecutive days. After
384 mesoderm induction, EBs are transferred to a 6-well plate with 20 EBs/well and placed in
385 macrophage precursor medium (X-vivo15 (Lonza, BE02-060Q), 100 ng/mL M-CSF
386 (Peprotech, 300-25), 25 ng/mL IL-3 (Peprotech, 213-13), 1x Glutamax, 50 U/ml Penicillin-
387 Streptomycin, 50 µM 2-Mercaptoethanol). 14 days after macrophage differentiation,
388 macrophage precursors were harvested using a cell strainer (Stemcell technologies, 27250).
389 Macrophage precursors were added to the NPC culture in 1:1 neural maintenance
390 medium:microglia medium ((DMEM-F12+Glutamax, 100 ng/mL M-CSF (Peprotech, 300-25),
391 100 ng/mL IL-34 (Peprotech, 200-34), 1x Glutamax, 50 U/ml Penicillin-Streptomycin, 50 µM
392 2-Mercaptoethanol).

393 **Replication labelling.** Prior to co-seeding of 1321N1 and SH-SY5Y mixed cultures, individual
394 cultures were incubated with either 10 µM EdU (Click-iT® EdU Imaging Kit, Life
395 Technologies, C10340) or 10 µM BrdU (Sigma-Aldrich, B5002) for 24h. This incubation time
396 exceeded the doubling time, allowing incorporation of the nucleotide analog in all cells. This
397 labelling period was followed by a 24h washout period in regular cell culture medium. After
398 washout, the cells were subcultured and plated in coculture. In half of the replicate, SH-SY5Y
399 cells received BrdU while 1321N1 cells received EdU. For the remainder of wells, the pre-
400 label switched cell types.

401 **Morphological staining.** Morphological staining (Cell Painting) was adapted from Bray et al.
402 2016[13]. After careful titration, all dye concentrations were adjusted and evaluated for
403 compatibility with 4-channel laser and filter combinations available on the confocal
404 microscope (see further). Staining was performed on cell cultures fixed in 4% PFA (roti-
405 histofix 4% paraformaldehyde, Roth, 3105.2) for 15min. Cells were rinsed once with PBS-/-
406 (Life Technologies, 10010015) prior to fixation and 3x 5min post fixation. Staining solutions

407 were prepared fresh before staining in PBS-/- with 0.3% Triton-X-100 (Sigma Aldrich, X100)
408 (**Table 1**). Each staining solution was incubated for 30min on a shaker at RT in the dark. After
409 staining, the cells were washed 1x with PBS -/- and sealed for imaging.

410 **Cyclic staining and immunocytochemistry.** Cyclic staining is executed by fluorescence
411 quenching after each sequential imaging round. 1mg/ml in ddH_2O $LiBH_4$ solution[22] (Acros
412 Organics, 206810050) was prepared fresh before use. 1.5h incubation of quenching solution
413 was performed before each successive staining series. After incubation, the quenching solution
414 was removed by washing 3x 5min in PBS-/-. Successful fluorescence quenching was
415 microscopically verified before proceeding with immunofluorescence staining (IF) (**Table 2**).
416 Cells are treated with PAV blocking buffer (Thimerosal 0.5% (Fluka, 71230), NaN3 0.1%
417 (Merck, k 6688), Bovine serum albumin (Sigma-Aldrich, A7284), Normal horse serum, PBS-
418 /-) for 8min. The desired primary antibodies (pAB) are diluted in PAV blocking buffer. pAB
419 incubation was performed 12h (overnight) at 4°C after which the cells were washed 1x 5min
420 in PBS-/- followed by incubation in secondary antibody solution (sAB) in PAV + DAPI for
421 nuclear counterstain. sAB staining was performed for 3h at RT while shaking. Prior to imaging,
422 the cells were washed 2x with PBS-/- and stored in PBS-/- + 0,1% NaN_3.

423 BrdU staining was performed using IF, requiring DNA denaturation before pAB incubation.
424 This was performed by 10 min incubation with 2N HCl at 37°C. HCl was neutralized with 0.1
425 M sodium borate buffer pH 8.5 for 30min at RT. Cells were washed 3x 5min in PBS-/- before
426 continuing with the general IF protocol. EdU click-it labelling was performed according to the
427 manufacturer's instructions (Click-iT® EdU Imaging Kit, Life Technologies, C10340).

428 **Image acquisition.** Images were acquired using a spinning disk confocal microscope (Nikon
429 CSU-W1 SoRa) with a 20x 0.75 NA objective (Plan APO VC 20x DIC N2) and Kinetix
430 sCMOS camera (pinhole 50µm; disk speed 4000 rpm; pinhole aperture 10; bit depth 12-bit,
431 pixel size 0.325µm²). We opted for confocal microscopy instead of widefield to overcome

432 image quality limitations resulting from highly dense cell clusters. 96-well plates were scanned,
433 capturing a minimum of 16 images per well spread in a regular pattern (0,8 mm spacing in x
434 and y) across the surface of the well. If multiple z-slices were acquired (to correct for surface
435 inclinations in the field of view), a maximum projection was performed before further analysis.
436 Morphological images were acquired in all 4 channels. (**Table 3**).
437 **Software.** Images were captured by the Nikon CSU-W1 confocal system in combination with
438 the NIS elements software (RRID:SCR_014329). Image visualization was later performed
439 using Fiji freeware[36,37]. In-depth image analysis, pre-processing and machine learning for cell
440 classification were performed using Python programming language[38] in combination with
441 Anaconda[39] (distribution and package managing software) and Visual Studio Code (code
442 editor). The packages and versions used for data analysis are shown in **Table 4**.
443 **Image pre-processing.** *Cell/nuclear segmentation.* Using a tailor-made script implementing
444 the Cellpose[40] (cell segmentation) or StarDist[41] (nuclear segmentation) package, images were
445 pre-processed by normalizing the intensity values of each channel between the 1st and 99th
446 quantile. Individual channels or channel combinations for segmentation can be selected
447 depending on the desired outcome mask. Resulting outputs included the segmentation mask
448 alongside the quality control images. For Stardist implementation, hyperparameter probability
449 was set at 0,6 and overlap at 0,3. For Cellpose segmentation, 4 models (cyto2) were averaged
450 to obtain the final mask. Segmentation was performed on the composite of all CP channels. An
451 estimation of the cell's diameter could be included to optimize cell detection.
452 *Ground truth alignment.* Following sequential staining and imaging rounds, multiple images
453 were captured representing the same cell with different markers. All images were aligned using
454 Fourier-based image cross correlation on the intensity-normalized multichannel images. The
455 alignment shift between image1 and image2 was determined using scikit-image phase cross

456 correlation. Image2 was then moved according to the predetermined shift to align
457 morphological with ground truth images.
458 *Ground truth phenotyping.* The true cell phenotype was determined by the fluorescence
459 intensity of the post-hoc immunostaining with class-specific markers. Each ground truth image
460 was imported alongside the corresponding cell mask for that image. For each cell label, the
461 fluorescence intensity was determined and tabulated. The threshold was set manually based on
462 the fluorescence intensity of the monoculture controls. Ground truth labels were assigned to
463 each region of interest (ROI).
464 *Data filtering.* Over the entire pipeline, ROIs could be discarded based on 3 conditions: (1)
465 Cell detachment. Due to the harsh sample preparation and repeated washing steps, cells could
466 detach from the imaging substrate thus resulting in lack of ground truth information for those
467 cells. As a result, all ROIs for which there was no DAPI signal detected in the ground truth
468 image, were removed from the dataset as incomplete. (2) Cells for which the ground truth
469 fluorescence intensity was ambiguous. Ground truth labels were determined based on the
470 presence of specific IF staining in either of the phenotype classes. If no class-specific IF
471 staining was detected by thresholding, no true label could be assigned. These ROIs were
472 therefore discarded due to uncertainty. (3) Likelihood of faulty ROI detection. The DAPI signal
473 was used to discard ROIs that do not represent (complete) cells by thresholding on minimal
474 DAPI intensity and minimal nuclear size.
475 **ROI classification.** *Train-validation-test split.* Model training requires splitting the available
476 data in 3 subsets: training (60%), validation (10%), and testing (30%). The training dataset was
477 used to train the machine learning models (either RF or CNN). The validation dataset,
478 composing of 10% of the total data, was used to determine the hyperparameters and
479 intermediate testing. The remaining 30% of the dataset was kept apart and used to test the final
480 model when training is completed. For both RF and CNN, the testing dataset was never shown

481 to the model during the training phase, but only used after training to determine the accuracy
482 of predictions to the ground truth. The number of instances for each class was equalized by
483 sampling equal number of instances from each predicted class. To account for variation
484 between technical replicates, the train-validation-test split was stratified per well. As a result,
485 no datapoints arising from the same well could appear simultaneously in the training, validation
486 and testing subset. This data stratification was repeated 3 times with different random seeds
487 (see Methods – Reproducibility).

488 *Random Forest.* For each ROI, a set of manually defined parameters was extracted
489 corresponding to cell shape (area, convex area, filled area, length minor axis, length major axis,
490 centroid, eccentricity, equivalent diameter area, ferret diameter max, orientation, perimeter),
491 texture (GLCM: contrast, dissimilarity, homogeneity, energy, correlation, ASM), intensity
492 (maximal intensity, minimal intensity, mean intensity, standard deviation intensity) and
493 distribution (solidity). This was done for 3 regions (nucleus, cytoplasm and complete cell), and
494 for all channel dimensions. Redundant parameters were removed. All parameters were
495 standardized per ROI, grouped per replicate. The number of trees within the forest was varied
496 between 10 and 150, reaching maximum accuracy at around 30 trees.

497 *Uniform Manifold Approximation and Projection.* Dimensionality reduction using UMAP was
498 performed on the same feature matrix as used for RF prediction. Hyperparameters were set at
499 the default settings.

500 *Convolutional neural network.* A ResNet50 model was trained for image classification. In
501 contrast to classical machine learning techniques, no handcrafted features were extracted. Cell
502 crops (60µm whole cell – 15µm nucleocentric/nuclear area) were defined based on the
503 segmentation mask for each ROI. Each ROI was cropped out of the original morphological
504 image and associated with metadata corresponding to its ground truth label. Images alongside
505 their labels were fed to the network. Tensors were normalized per channel. Models are trained

506 on a minimum of 5000 training inputs of each class for 50 epochs (training iterations). Each
507 batch consisted of 100 samples. The training input was augmented by linear transformations
508 (horizontal and vertical flip, random rotation). Each epoch, the current model was tested against
509 a validation dataset. The performance of the model on this validation subset determined
510 whether the model was stored as new best (if the new accuracy exceeded the accuracy of the
511 previous best model) or discarded. The learning rate at the start was set at 0,0001 and
512 automatically reduced with a factor of 0,1 during training when no improvement was seen after
513 10 epochs. After 50 epochs, the best resulting model was tested on a separate test dataset to
514 determine the accuracy on previously unseen data.

515 **Reproducibility.** Each model training was performed 3 independent times. This was repeated
516 for 3 different random seeds. Each model received input data arising from a minimum of 16
517 images per well, at least 15 technical replicates (wells). The optimization experiments (figures
518 1-4) were performed with cell lines with limited variability. These models were trained on 3
519 independent experiments where ground truth pre-labelling (Edu/BrdU) was performed at least
520 once on either of the cell lines in coculture. For iPSC-derived cultures, as variability is inherent
521 to these differentiations, 3 biological replicates (independent differentiations) were pooled for
522 model training.

523 **Statistics**. All statistical comparisons were made nonparametric using Mann-Whitney U (for
524 two independent sample comparison) or Kruskal-Wallis (for multiple sample comparison) with
525 pairwise tests using Tukey's honestly significant difference test. We opted for nonparametric
526 testing because the number of models in each group to be compared was < 15. Significance
527 levels are indicated on the figures using ns. (no statistical significance, p-value above 0,05), *
528 (p-value between 0,05 and 5e-4), ** (p-value between 5e-4 and 5e-6) and *** (p-value smaller
529 than 5e-6).

530

Funding

This work was funded by Fonds Wetenschappelijk Onderzoek Vlaanderen (1SB7423N; 1274822N), BOF (FFB210009) and IOF Uantwerpen (FFI210239; FFI230099) and VLAIO (HBC.2023.0155).

Acknowledgements

We thank Marlies Verschuuren and Hugo Steenberghen for their assistance and knowledge regarding the cell lines. By extension, we would like to thank all current and former members of the De Vos lab.

Data availability statement

The authors report that the results of this study are available within the manuscript and supplementary materials. All image analysis scripts are open-source available on GitHub (https://github.com/DeVosLab)

Conflict of interest

The authors declare no conflict of interest.

Abbreviations

BrdU: Bromodeoxyuridine; CNN: convolutional neural network; CP: cell painting; DAPI: 4',6-diamidino-2-fenylindool; DIV: days in vitro; EdU: 5-ethynyl-2'-deoxyuridine; ER: endoplasmic reticulum; Grad-CAM: Gradient-weighted Class Activation Mapping; IF:

553 immunofluorescence; iPSC: induced pluripotent stem cell; NPC: neural progenitor cell; RF:
554 random forest; ROI: region of interest; TUBB3: beta-III-tubulin; UMAP: Uniform Manifold
555 Approximation and Projection

556

557 **Author contributions**

558 The experiments were conceptualized by SDB, TVDL, JVDD, PP and WDV. Experiments
559 were executed by SDB and JVDD. TVDL and SDB optimized the data analysis scripts. SDB
560 analysed the imaging data and performed the data analysis. SDB and WDV prepared the
561 original draft. PP and WDV supervised the work. All authors took part in reviewing and editing
562 the manuscript.

563

References

1. De Strooper, B. & Karran, E. The Cellular Phase of Alzheimer's Disease. *Cell* **164**, 603–615 (2016).

2. Bell, S. *et al.* Differentiation of Human Induced Pluripotent Stem Cells (iPSCs) into an Effective Model of Forebrain Neural Progenitor Cells and Mature Neurons. *BIO-PROTOCOL* **9**, (2019).

3. Neyrinck, K. *et al.* SOX9-induced Generation of Functional Astrocytes Supporting Neuronal Maturation in an All-human System. *Stem Cell Rev and Rep* **17**, 1855–1873 (2021).

4. Haenseler, W. *et al.* A Highly Efficient Human Pluripotent Stem Cell Microglia Model Displays a Neuronal-Co-culture-Specific Expression Profile and Inflammatory Response. *Stem Cell Reports* **8**, 1727–1742 (2017).

5. Nevin, Z. S. *et al.* Modeling the Mutational and Phenotypic Landscapes of Pelizaeus-Merzbacher Disease with Human iPSC-Derived Oligodendrocytes. *The American Journal of Human Genetics* **100**, 617–634 (2017).

6. Abutaleb, N. O. & Truskey, G. A. Differentiation and characterization of human iPSC-derived vascular endothelial cells under physiological shear stress. *STAR Protocols* **2**, 100394 (2021).

7. Paik, D. T., Chandy, M. & Wu, J. C. Patient and Disease–Specific Induced Pluripotent Stem Cells for Discovery of Personalized Cardiovascular Drugs and Therapeutics. *Pharmacol Rev* **72**, 320–342 (2020).

8. Lopez-Gonzalez, R. *et al.* Poly(GR) in C9ORF72 -Related ALS/FTD Compromises Mitochondrial Function and Increases Oxidative Stress and DNA Damage in iPSC-Derived Motor Neurons. *Neuron* **92**, 383–391 (2016).

9. Amin, N. D. & Paşca, S. P. Building Models of Brain Disorders with Three-Dimensional Organoids. *Neuron* **100**, 389–405 (2018).

10. D'Antonio, M. *et al.* High-Throughput and Cost-Effective Characterization of Induced Pluripotent Stem Cells. *Stem Cell Reports* **8**, 1101–1111 (2017).

11. Chen, C. X.-Q. *et al. Standardized quality control workflow to evaluate the reproducibility and differentiation potential of human iPSCs into neurons*. http://biorxiv.org/lookup/doi/10.1101/2021.01.13.426620 (2021) doi:10.1101/2021.01.13.426620.

12. Gustafsdottir, S. M. *et al.* Multiplex Cytological Profiling Assay to Measure Diverse Cellular States. *PLoS ONE* **8**, e80999 (2013).

13. Bray, M.-A. *et al.* Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. *Nat Protoc* **11**, 1757–1774 (2016).

14. Cimini, B. A. *et al.* Optimizing the Cell Painting assay for image-based profiling. *Nat Protoc* **18**, 1981–2013 (2023).

15. Way, G. P. *et al.* Predicting cell health phenotypes using image-based morphology profiling. *MBoC* **32**, 995–1005 (2021).

16. Schiff, L. *et al.* Integrating deep learning and unbiased automated high-content screening to identify complex disease signatures in human fibroblasts. *Nat Commun* **13**, 1590 (2022).

17. Chandrasekaran, S. N. *et al. JUMP Cell Painting dataset: morphological impact of 136,000 chemical and genetic perturbations*. http://biorxiv.org/lookup/doi/10.1101/2023.03.23.534023 (2023) doi:10.1101/2023.03.23.534023.

18. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. (2018) doi:10.48550/ARXIV.1802.03426.

19. Nguyen, T.-T., Huang, J. Z. & Nguyen, T. T. Unbiased Feature Selection in Learning Random Forests for High-Dimensional Data. *The Scientific World Journal* **2015**, 1–18 (2015).

20. He, K., Zhang, X., Ren, S. & Sun, J. Deep Residual Learning for Image Recognition. Preprint at http://arxiv.org/abs/1512.03385 (2015).

21. Selvaraju, R. R. *et al.* Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *Int J Comput Vis* **128**, 336–359 (2020).

22. Radtke, A. J. *et al.* IBEX: an iterative immunolabeling and chemical bleaching method for high-content imaging of diverse tissues. *Nat Protoc* **17**, 378–401 (2022).

23. Hernández, D. *et al.* Culture Variabilities of Human iPSC-Derived Cerebral Organoids Are a Major Issue for the Modelling of Phenotypes Observed in Alzheimer's Disease. *Stem Cell Rev and Rep* **18**, 718–731 (2022).

24. Volpato, V. & Webber, C. Addressing variability in iPSC-derived models of human disease: guidelines to promote reproducibility. *Disease Models & Mechanisms* **13**, dmm042317 (2020).

25. Shi, Y., Kirwan, P. & Livesey, F. J. Directed differentiation of human pluripotent stem cells to cerebral cortex neurons and neural networks. *Nat Protoc* **7**, 1836–1846 (2012).

26. Kuijlaars, J. *et al.* Sustained synchronized neuronal network activity in a human astrocyte co-culture system. *Sci Rep* **6**, 36529 (2016).

27. Dincer, A. B., Celik, S., Hiranuma, N. & Lee, S.-I. *DeepProfile: Deep learning of cancer molecular profiles for precision medicine*. http://biorxiv.org/lookup/doi/10.1101/278739 (2018) doi:10.1101/278739.

28. De Vos, W. H. *et al.* Increased plasticity of the nuclear envelope and hypermobility of telomeres due to the loss of A–type lamins. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1800**, 448–458 (2010).

29. Heckenbach, I. *et al.* Nuclear morphology is a deep learning biomarker of cellular senescence. *Nat Aging* **2**, 742–755 (2022).

30. Mikolajczyk, A. & Grochowski, M. Data augmentation for improving deep learning in image classification problem. in *2018 International Interdisciplinary PhD Workshop (IIPhDW)* 117–122 (IEEE, 2018). doi:10.1109/IIPHDW.2018.8388338.

31. Nguyen, L. D., Lin, D., Lin, Z. & Cao, J. Deep CNNs for microscopic image classification by exploiting transfer learning and feature concatenation. in *2018 IEEE International Symposium on Circuits and Systems (ISCAS)* 1–5 (IEEE, 2018). doi:10.1109/ISCAS.2018.8351550.

32. Tomov, M. L. *et al.* Resolving cell state in iPSC-derived human neural samples with multiplexed fluorescence imaging. *Commun Biol* **4**, 786 (2021).

33. Strano, A., Tuck, E., Stubbs, V. E. & Livesey, F. J. Variable Outcomes in Neural Differentiation of Human PSCs Arise from Intrinsic Differences in Developmental Signaling Pathways. *Cell Reports* **31**, 107732 (2020).

34. Galiakberova, A. A. *et al.* Different iPSC-derived neural stem cells shows various spectrums of spontaneous differentiation during long term cultivation. *Front. Mol. Neurosci.* **16**, 1037902 (2023).

35. Du, J. & Xu, Y. Hierarchical deep neural network for multivariate regression. *Pattern Recognition* **63**, 149–157 (2017).

36. Rueden, C. T. *et al.* ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics* **18**, 529 (2017).

37. Schindelin, J., Rueden, C. T., Hiner, M. C. & Eliceiri, K. W. The ImageJ ecosystem: An open platform for biomedical image analysis. *Molecular Reproduction Devel* **82**, 518–529 (2015).

38. Van Rossum, G., & Drake, F. L. Python 3 Reference Manual. (2009).

39. Anaconda Software Distribution. Anaconda Documentation.

40. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nat Methods* **18**, 100–106 (2021).

41. Weigert, M., Schmidt, U., Haase, R., Sugawara, K. & Myers, G. Star-convex Polyhedra for 3D Object Detection and Segmentation in Microscopy. in *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)* 3655–3662 (IEEE, 2020). doi:10.1109/WACV45572.2020.9093435.

653 **Tables**

Table 1 | Specifications of morphological staining composition.

Staining solution 1							
Dye	**Target**	**Excitation (nm)**	**Emission (nm)**	**Stock concentration**	**End dilution**	**Supplier**	**Catalog #**
DAPI	Nucleus	350	470	2.5 mg/ml	1/500	Sigma Aldrich	D9542
Concanavalin	ER	480	490-540	5 mg/ml	1/300	Life Technologies	C11252
WGA	Golgi	550	560-570	1 mg/ml	1/600	Life Technologies	W32464
Phalloidin	Actin	570	580-610	66µM	1/1000	Life Technologies	A12380
Mitotracker	Mitochondria	640	650-660	1 mM	1/1000	Life Technologies	M22426
Staining solution 2							
Syto14	Nucleoli/RNA	525	540-580	5 mM	1/1000	Life Technologies	S7576

654

Table 2 | Used antibodies.

Antibody		Host	Supplier	Catalog #
Primary antibodies	Anti-BrdU	Sheep	Abcam	ab1893
	Anti-MAP2	Chicken	Synaptic systems	188006
	Anti-TUBB3	Mouse	BioLegend	801202
	Anti-CD44	Mouse	Millipore	MAB4065
	Anti-GFAP	Goat	Abcam	ab53554
	Anti-NeuN	Guinea pig	Millipore	ABN90P
	Anti-Iba1	Rabbit	Wako	019-19741
	Anti-Ki67	Rabbit	Abcam	ab15580
Secondary antibodies	Anti-mouse-Cy3	Donkey	Jackson ImmunoResearch	715-165-150
	Anti-mouse-Cy5	Donkey	Jackson ImmunoResearch	715-175-150
	Anti-sheep-Cy3	Donkey	Jackson ImmunoResearch	713-165-147
	Anti-sheep-Cy5	Donkey	Jackson ImmunoResearch	713-175-147
	Anti-chicken-Cy5	Donkey	Jackson ImmunoResearch	703-175-155
	Anti-goat-Cy3	Donkey	Jackson ImmunoResearch	705-165-147
	Anti-Guinea pig-Cy3	Donkey	Jackson ImmunoResearch	706-165-148
	Anti-Rabbit-Cy5	Donkey	Jackson ImmunoResearch	711-175-152

655

Table 3 | Specifications of the used laser lines, excitation and emission filters.

Channel dimension	Laser line (nm)	Excitation filter (nm)	Emission filter (nm)

Channel 1	405	Di01-T405/488/568/647-13x15x0.5	B447/60
Channel 2	488		B520/35
Channel 3	561		B617/73
Channel 4	640		B685/40

656

Table 4 | Python packages used for image and data analysis alongside the software version.

Package	**Version**	**Reference**
pathlib	1.0.1	
numpy	1.21.2	Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., … Oliphant, T. E. (2020). Array programming with NumPy. Nature, 585, 357–362. https://doi.org/10.1038/s41586-020-2649-2
tifffile	2021.7.2	
pytorch	1.13.0	Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., … Chintala, S. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. In Advances in Neural Information Processing Systems 32 (pp. 8024–8035). Curran Associates, Inc.
torchvision	0.14.0	
nd2reader	3.3.0	
matplotlib	3.5.2	Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. Computing in Science & Engineering, 9(3), 90–95.
scikit-image	0.19.3	Van der Walt, S., Sch"onberger, Johannes L, Nunez-Iglesias, J., Boulogne, Franccois, Warner, J. D., Yager, N., … Yu, T. (2014). scikit-image: image processing in Python. PeerJ, 2, e453.

scikit-learn	1.1.3	Pedregosa, F., Varoquaux, Ga"el, Gramfort, A., Michel, V., Thirion, B., Grisel, O., … others. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12(Oct), 2825–2830.
argparse	1.4.0	
cellpose	2.1.1	Stringer, C., Wang, T., Michaelos, M. et al. Cellpose: a generalist algorithm for cellular segmentation. Nat Methods 18, 100–106 (2021). https://doi.org/10.1038/s41592-020-01018-x
tqdm	4.64.1	
scipy	1.9.3	Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., … SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods, 17, 261–272. https://doi.org/10.1038/s41592-019-0686-2
pandas	1.5.1	McKinney, W., & others. (2010). Data structures for statistical computing in python. In Proceedings of the 9th Python in Science Conference (Vol. 445, pp. 51–56).
seaborn	0.12.1	https://doi.org/10.5281/zenodo.883859
imbalanced-learn	0.10.0	Imbalanced-learn: a python toolbox to tackle the curse of imbalanced datasets in machine learning
StarDist	0.8.5	M. Weigert, U. Schmidt, R. Haase, K. Sugawara and G. Myers, "Star-convex Polyhedra for 3D Object Detection and Segmentation in Microscopy," 2020 IEEE Winter Conference on Applications of Computer Vision (WACV), Snowmass, CO, USA, 2020, pp. 3655-3662, doi: 10.1109/WACV45572.2020.9093435.

657

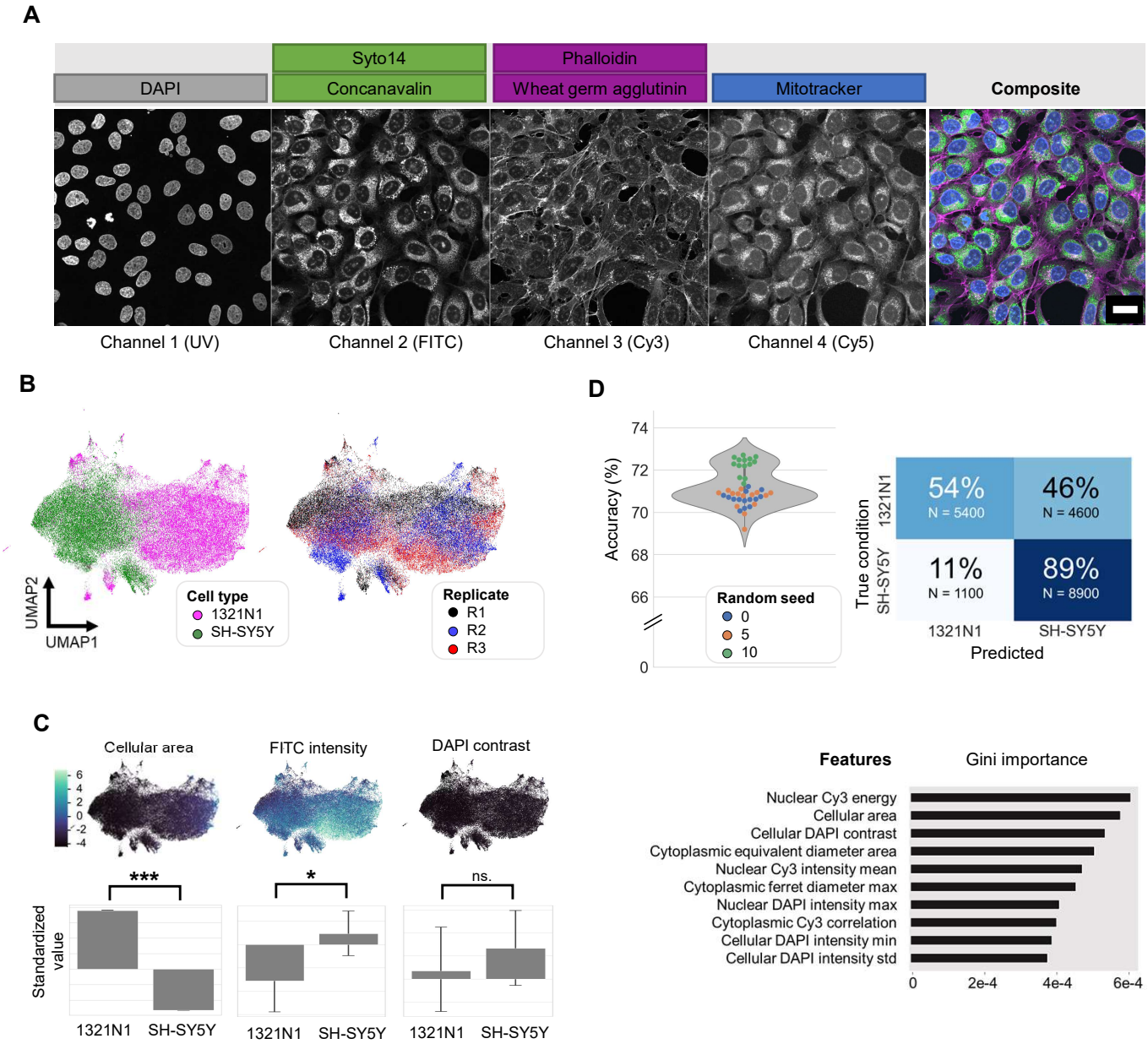


Figure 1 | Shallow classification using manually defined features. (A) Image overview of 1321N1 cells after CP staining acquired at higher magnification (Plan Apo VC 60xA WI DIC N2, NA = 1,20). Scale bar is 30µm. (B) UMAP dimensionality reduction, color coded for either cell type condition (left) or replicate (right). (C) Feature importance deducted from the UMAP (feature maps) and RF (Gini importance). Three exemplary feature maps are highlighted alongside the quantification per cell type. (D) Random Forest classification performance with confusion matrix and Gini importance. All features used in the UMAP are used for RF building.

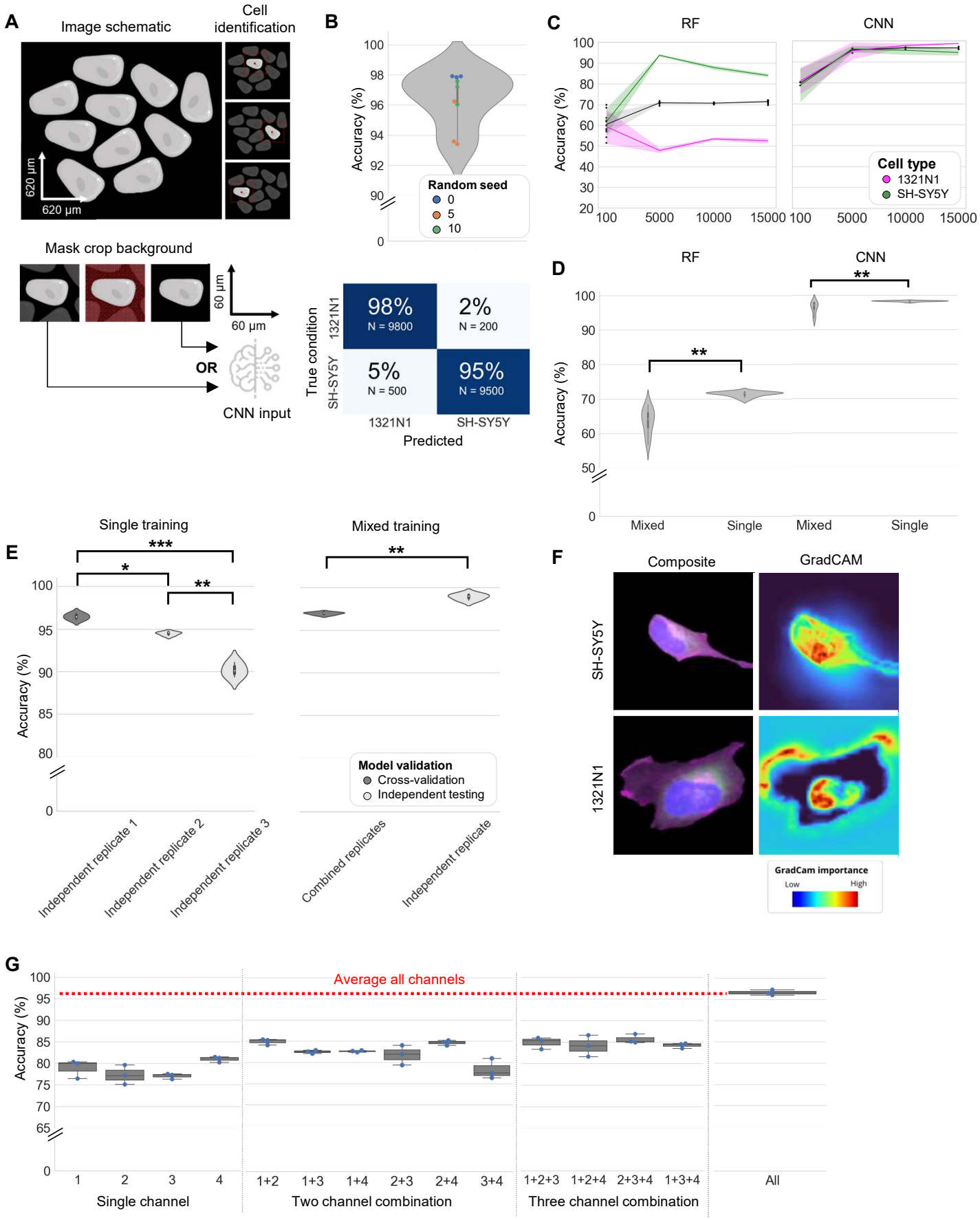


Figure 2 | Convolutional neural network classification of monoculture conditions. (A) Schematic of image pre-processing for CNN classification. (B) CNN accuracy and confusion matrix on cell type classification in monocultures. (C) Shallow vs. deep learning accuracy for varying training set sizes. Dots represent different training iterations (cross-validation). Train-validation-test dataset selection was performed with the same random seed at 60-10-30 ratios. (D) The impact of variability in the training dataset on model accuracy for shallow vs. deep learning. Mann-Whitney-U-test, p-values resp. 7,47e-6 and 1,96e-4. (E) The performance of deep learning classifiers trained on single replicates (low variability) and mixed replicates (high variability) on unseen images (independent replicates). Kruskal-Wallis test on single training condition, p-value of 0,026 with post-hoc Tukey test. Mann-Whitney-U-test on mixed training, p-value of 7,47e-6. (F) GradCAM map highlighting the regions on which the CNN relies for making its prediction. (G) Performance accuracy of the CNN when only a selection of image channels is given as training input.

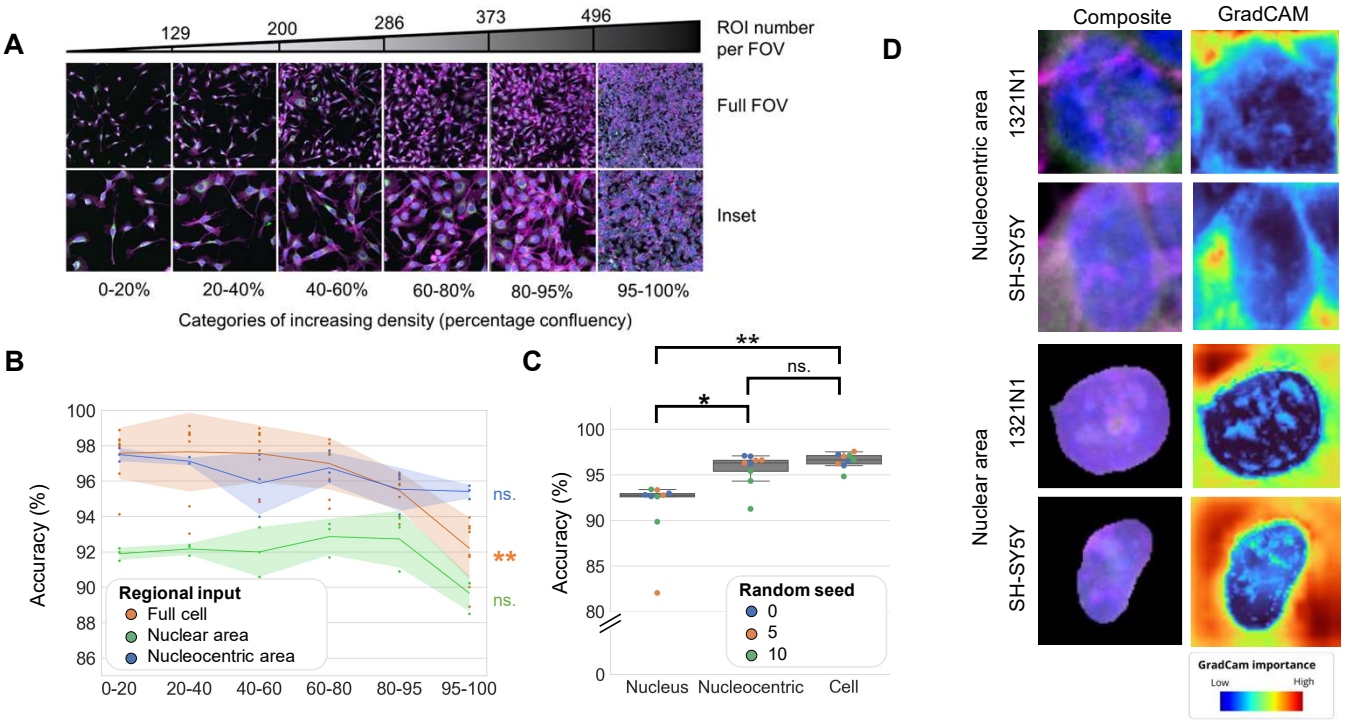


Figure 3 | The impact of regional information on CNN performance in relation to culture density. (A) Definition of different culture density categories with corresponding images. (B) The effect of culture density (number of cells within the FOV) on prediction accuracy, for each of the different input regions. Kruskal-Wallis test for each of the regions, statistically significant decrease (p-value 2,07e-5) between the most dense condition with all other density categories for the cell region. (C) CNN accuracy for different input regions. Kruskal-Wallis test, p-value of 4,29e-4 with post-hoc Tukey test. (D) GradCAM intensity plots highlighting the regions on which the CNN relies for making its prediction in nucleocentric (top) and nuclear (bottom) crops.

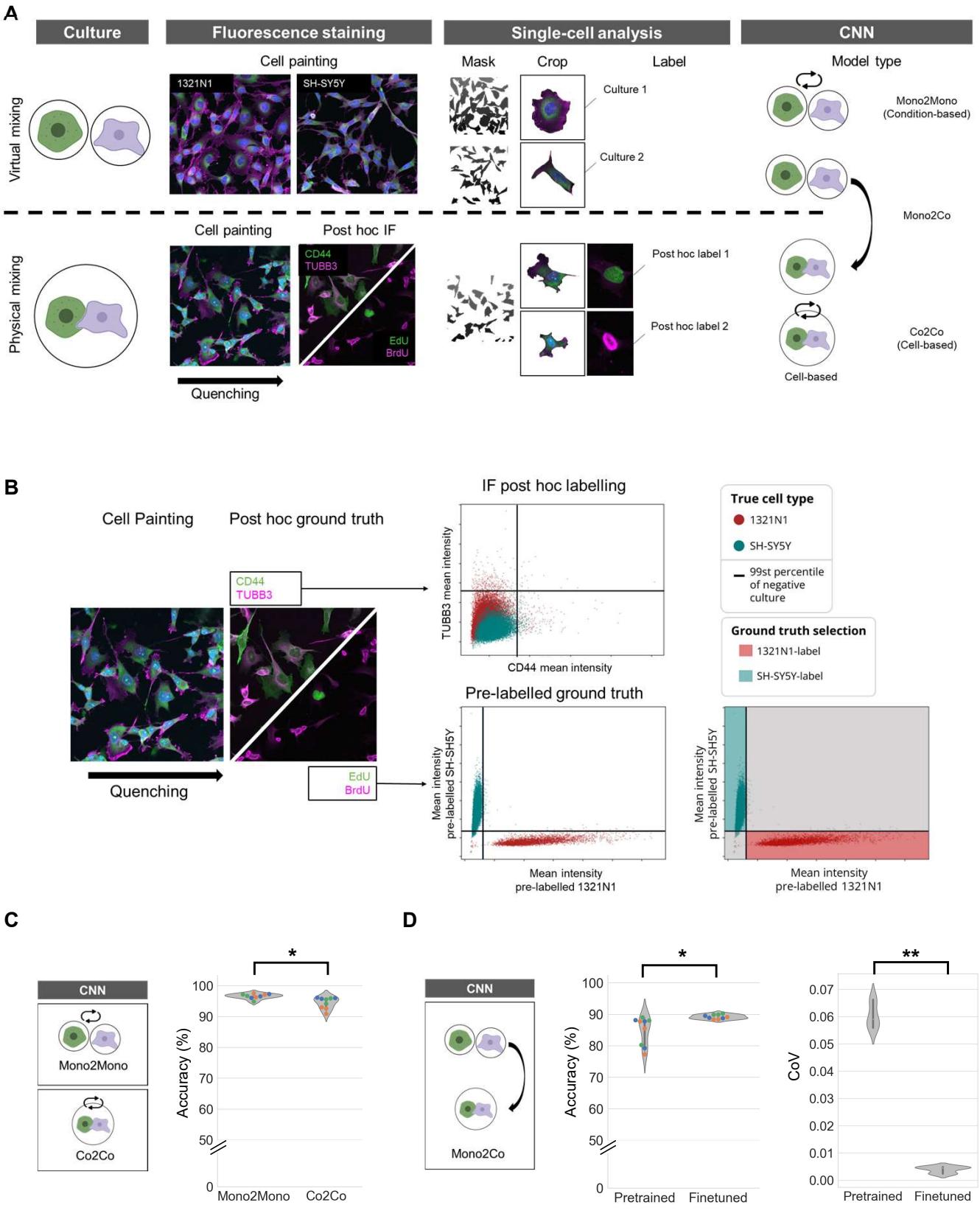


Figure 4 | Cell type prediction in mixed cultures by post-hoc ground truth identification. (A) Schematic overview of virtual vs. physically mixed cultures and the subsequent CNN models. For virtual mixing, individual cell types arise from distinct cultures, the cell type label was assigned based on the culture condition. For physical mixing, the true phenotype of each crop was determined by intensity thresholding after post-hoc staining. Three model types are defined: Mono2Mono (culture-based label), Co2Co (cell-based label) and Mono2Co (trained with culture-based label and tested on cell-based label). (B) Determining ground-truth phenotype by intensity thresholding on IF (top) or pre-label (bottom). The threshold was determined on the intensity of monocultures. Only the pre-labelled ground-truth resulted in good cell type separation by thresholding for 1321N1 vs. SH-SY5Y cells. (C) Mono2Mono (culture-based ground truth) vs. Co2Co (cell-based ground truth) models for cell type classification. Mann-Whitney-U-test p-value 0,0027. (D) Monoculture-trained models were tested on mixed cultures. Pretrained models were trained on independent biological replicates. These could be finetuned by additional training on monoculture images from the same replicate as the coculture. This was shown to reduce the variation between model iterations (Median performance: Mann-Whitney-U-test, p-value 0,0015; Coefficient of variation: Mann-Whitney-U-test, p-value 3,48e-4)

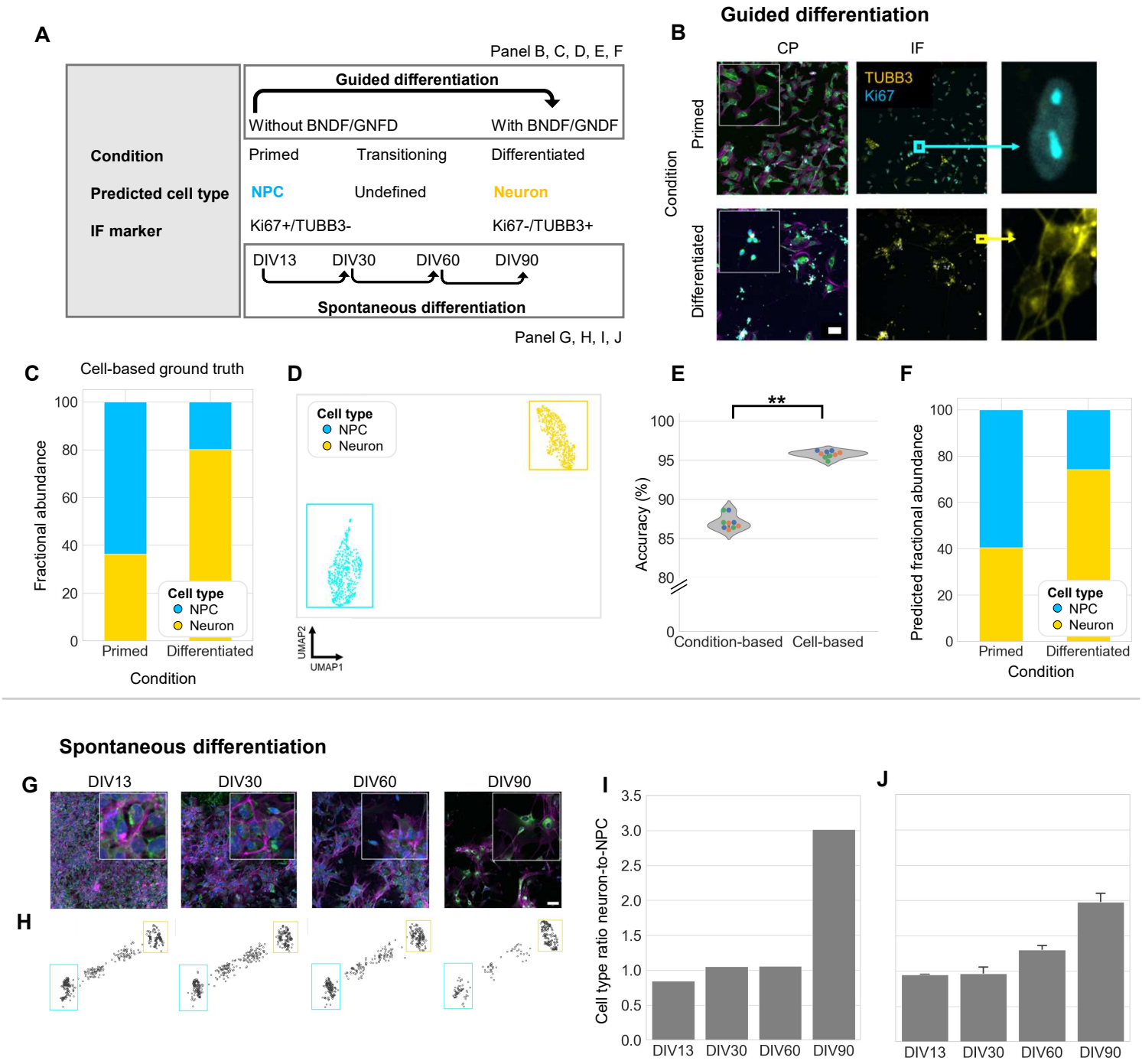


Figure 5 | iPSC-derived differentiation staging using morphology profiling. (A) Schematic overview of guided vs. spontaneous neural differentiation. (B) Representative images of morphological staining and post-hoc IF of primed and differentiated iPSC-derived cultures (guided differentiation). Ground-truth determination is performed using TUBB3 (for mature neurons) and Ki67 (mitotic progenitors). Scale bar is 200µm. (C) Fraction of neurons vs. NPC cells in the primed vs. differentiated condition as determined by IF staining. Upon guided differentiation, the fraction of neurons increased. (D) Following feature extraction (handcrafted features), supervised UMAP dimension reduction shows separation between the primed vs. differentiated cultures. (E) CNN performance when classifying neurons (Ki67-/TUBB3+) vs. NPC (Ki67+/TUBB3-) cells using either a condition-based or cell-type based ground truth. Mann-Whitney-U-test, p-value 4,04e-4. (F) Fractional abundance of predicted cell phenotypes (NPC vs. neurons) in primed vs. differentiated culture conditions using the cell-based CNN. (G) Representative images of spontaneously differentiating neural cultures. Scale bar is 200µm. (H) Handcrafted features were extracted from the morphological images. These morphological profiles were plotted onto the UMAP feature space created using the guided differentiation (panel D). Plotting the points of spontaneously differentiating cultures onto this space, revealed gradual displacement of the morphological profile from the NPC towards the neuronal cluster with increasing differentiation time. (I) Quantification of the number of points in each cluster of panel H. (J) Prediction of differentiation status using the cell-based CNN model trained on guided differentiated culture.

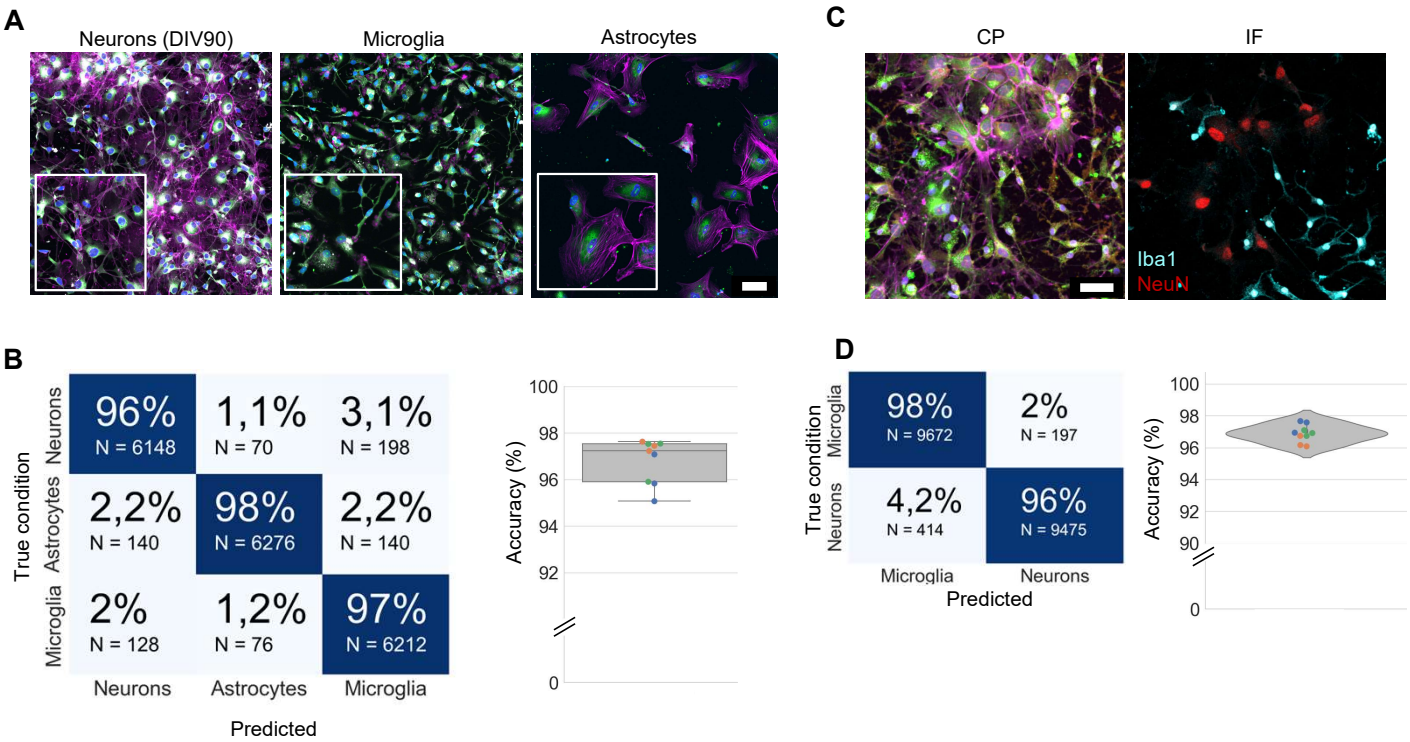
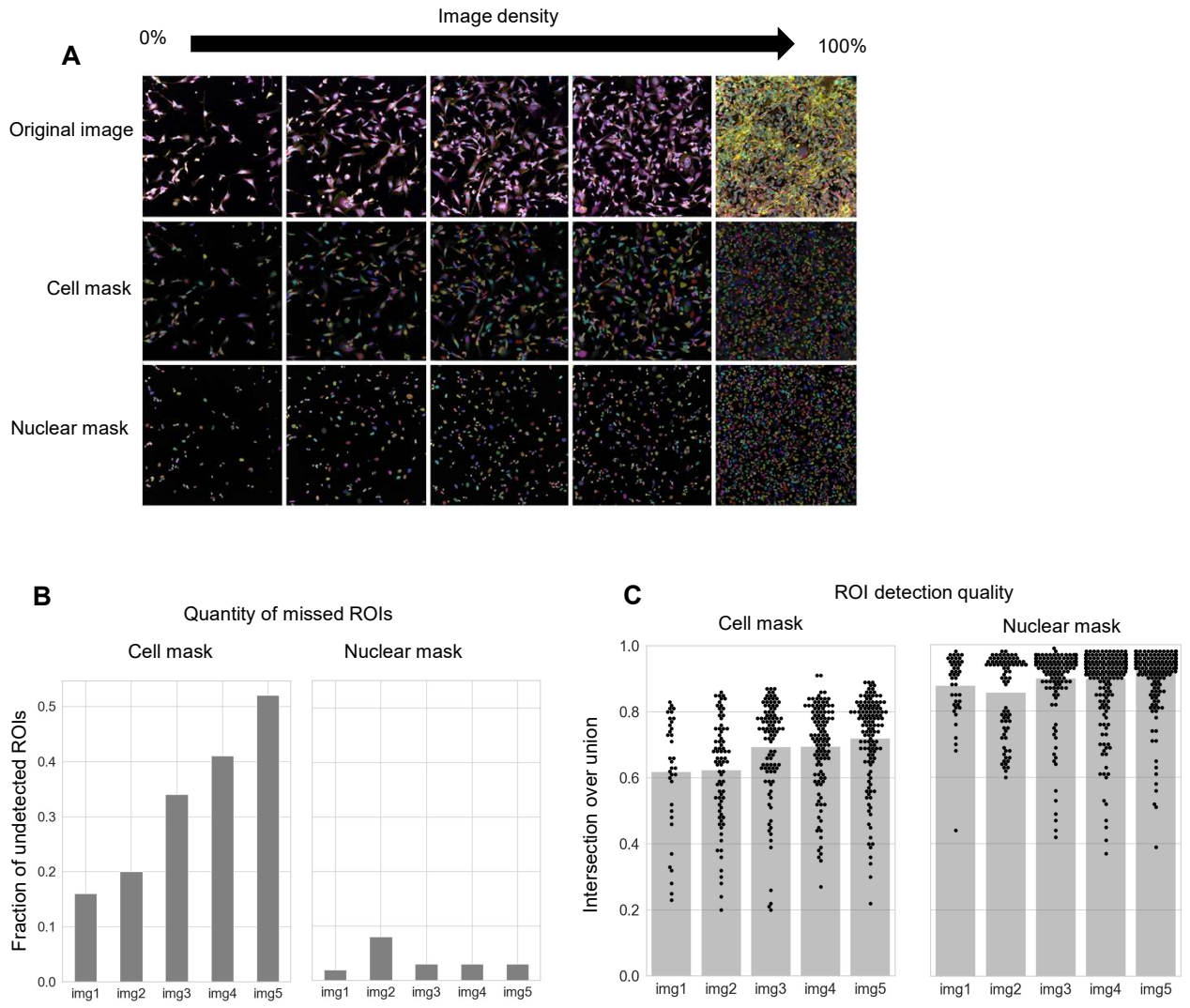
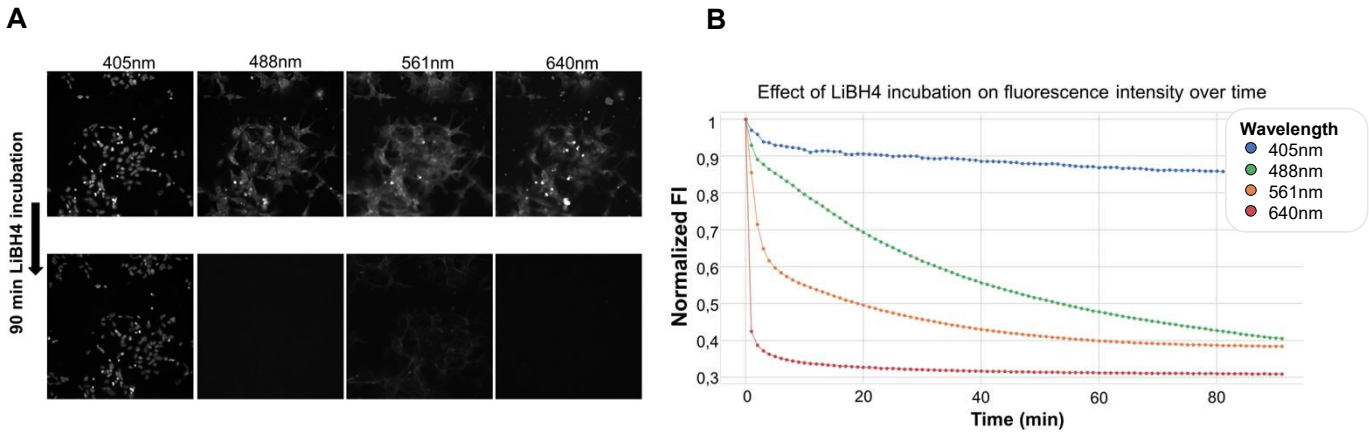


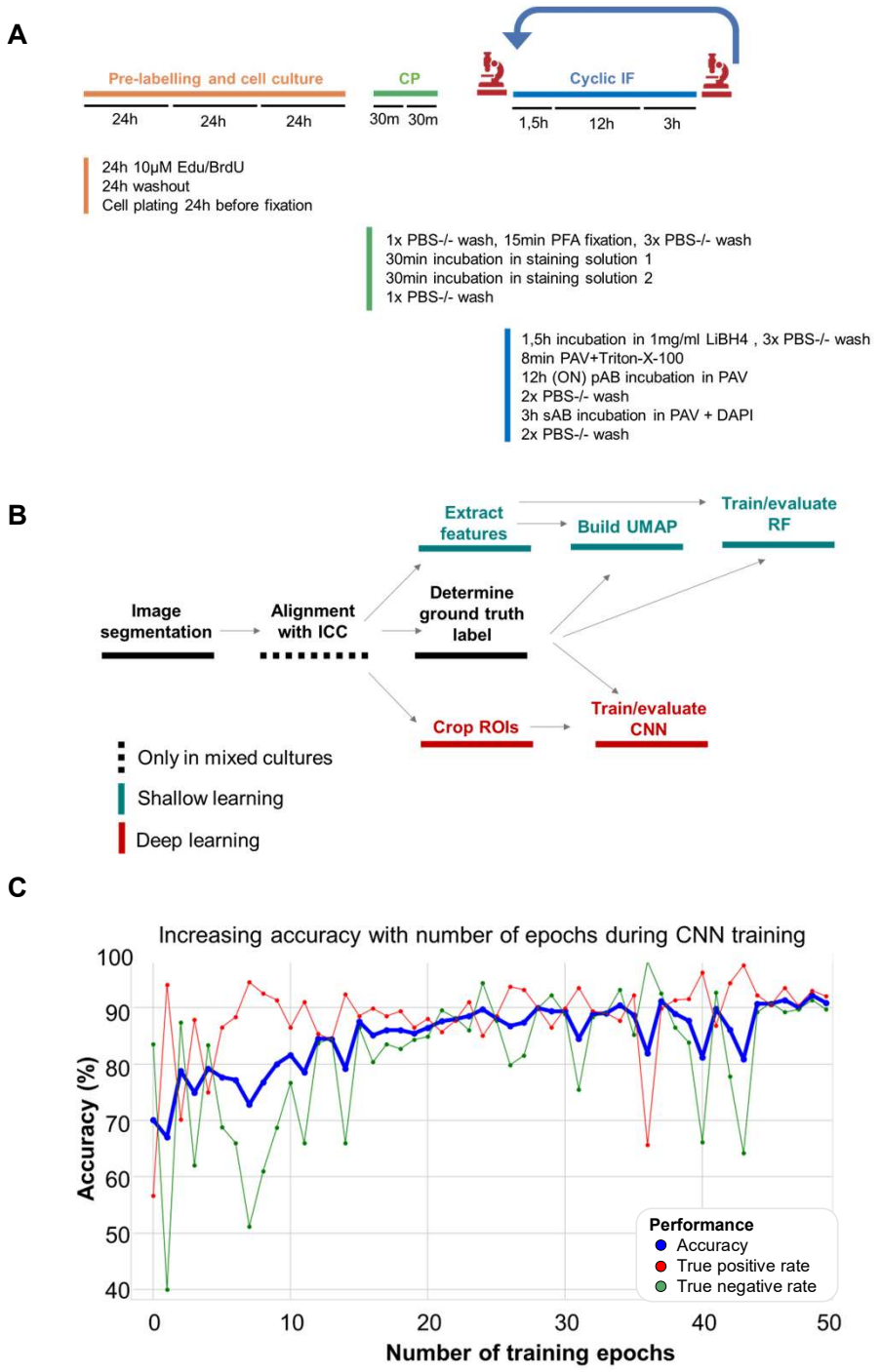
Figure 6 | iPSC cell type identification using morphology profiling. (A) Representative images of iPSC-derived neurons, astrocytes and microglia with morphological staining. Scale bar is 200µm. (B) Prediction accuracy of a CNN trained to classify monocultures of iPSC-derived astrocytes, microglia and neurons (panel A) with confusion matrix (average of all models). (C) Representative images of a mixed culture of iPSC-derived neurons and microglia. Ground-truth identification was performed using IF. (D) CNN accuracy of a model trained to identify both cell types in mixed culture mounted up to 98%.



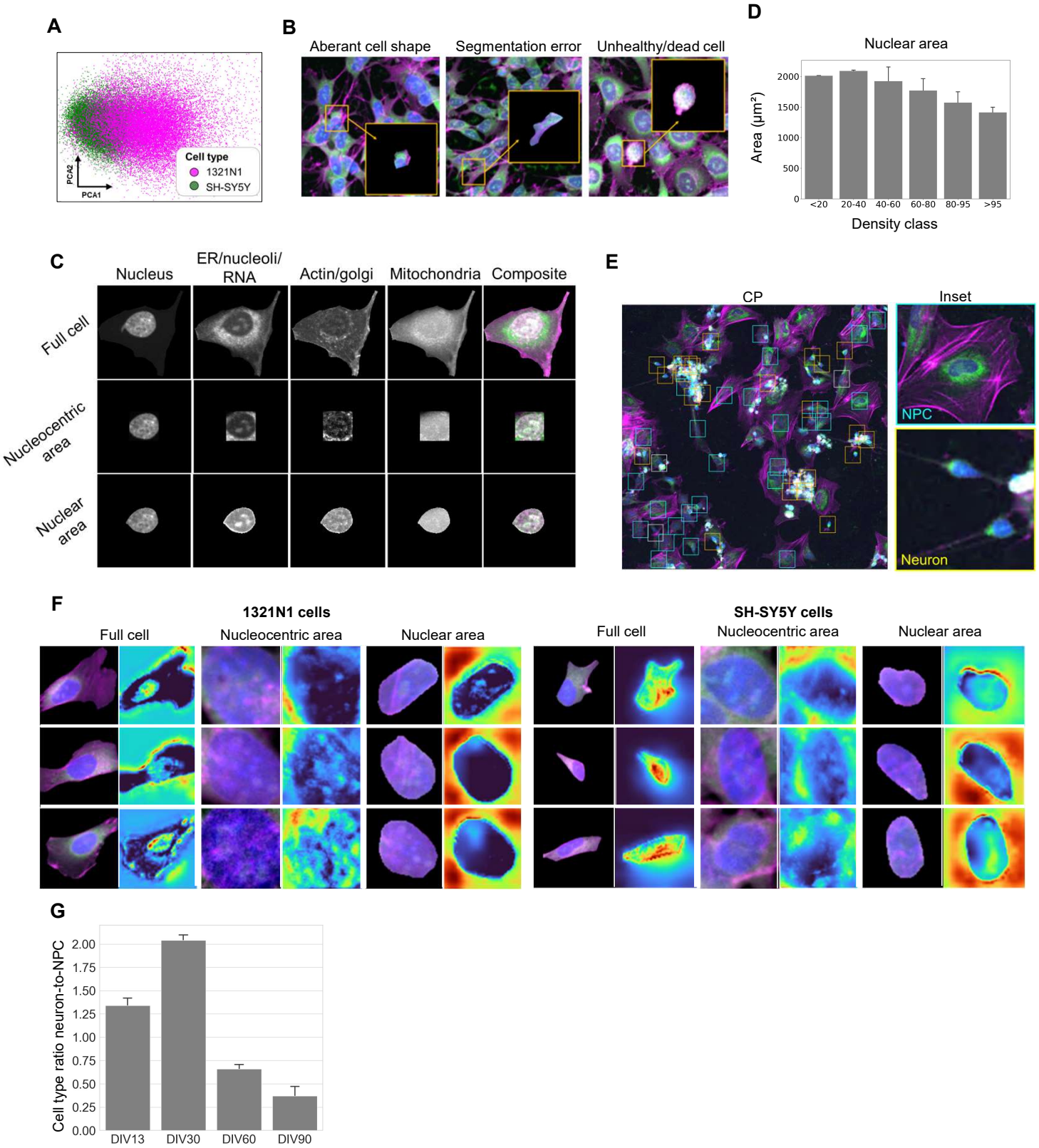
Supplementary figure 1 | The influence of cell density on segmentation quantity and quality. (A) Representative images of 1321N1 cells with increasing density alongside their cell and nuclear mask produced using resp. Cellpose and Stardist. (B) the number of ROIs detected in comparison to the ground truth (manual segmentation) (determined as intersection over union (IoU) < 0,15). (C) ROI detection quality (IoU) for increasing cell density for cell and nuclear masks.



Supplementary figure 2 | Fluorescence quenching over time using LiBH4. (A) Images before and after quenching for all 4 fluorescence channels. (B) Time curve of normalized image-level fluorescence intensity during incubation with 1mg/ml LiBH4.



Supplementary figure 3 | Methods. (A) Pipeline used for morphological profiling in mixed cultures. (B) Overview of the steps within the image analysis pipeline. (C) Evaluation of accuracy, true negative and true positive rate during CNN training (1321N1 vs. SH-SY5Y in monocultures) across all 50 epochs.



Supplementary figure 4 | Additional figures. (A) PCA dimension reduction on handcrafted features extracted from monocultures of 1321N1 and SH-SY5Y cells. In contrast to UMAP dimension reduction, this linear approach does not result in clear cell type clustering. (B) Examples of misclassified ROIs. (C) Definition of cell regions given as training input for nuclear and nucleocentric model training. (D) Decreasing nuclear area in function of culture density. (E) Evaluation of cell phenotype prediction on an image of differentiating neurons (guided differentiation). The predicted phenotype is shown by the color of the square around each ROI. The distinctive phenotype is shown in the insets. (F) Additional GradCAM images for both cell type classes. (G) Predicted neuron-to-NPC ratio in spontaneously differentiating cultures as predicted by the condition-based classifier (Fig. 5E) in contrast to the cell-based classifier performance on the same dataset (Fig. 5H).