



OPEN

## Label-free prediction of cell painting from brightfield images

Jan Oscar Cross-Zamirski<sup>1,2,5</sup>, Elizabeth Mouchet<sup>3</sup>, Guy Williams<sup>3</sup>, Carola-Bibiane Schönlieb<sup>1</sup>, Riku Turkki<sup>4,6</sup> & Yin Hai Wang<sup>2,6</sup>

Cell Painting is a high-content image-based assay applied in drug discovery to predict bioactivity, assess toxicity and understand mechanisms of action of chemical and genetic perturbations. We investigate label-free Cell Painting by predicting the five fluorescent Cell Painting channels from brightfield input. We train and validate two deep learning models with a dataset representing 17 batches, and we evaluate on batches treated with compounds from a phenotypic set. The mean Pearson correlation coefficient of the predicted images across all channels is 0.84. Without incorporating features into the model training, we achieved a mean correlation of 0.45 with ground truth features extracted using a segmentation-based feature extraction pipeline. Additionally, we identified 30 features which correlated greater than 0.8 to the ground truth. Toxicity analysis on the label-free Cell Painting resulted a sensitivity of 62.5% and specificity of 99.3% on images from unseen batches. We provide a breakdown of the feature profiles by channel and feature type to understand the potential and limitations of label-free morphological profiling. We demonstrate that label-free Cell Painting has the potential to be used for downstream analyses and could allow for repurposing imaging channels for other non-generic fluorescent stains of more targeted biological interest.

Cell Painting<sup>1</sup> is a high-content assay used for image-based profiling in drug discovery<sup>2</sup>. It is an inexpensive methodology that captures the rich information in cell morphology and has shown promising utility in bioactivity prediction, identification of cytotoxicity and prediction of mechanisms of action<sup>3–5</sup>. Cell phenotypes are captured with six generic fluorescent dyes and imaged across five channels, visualizing eight cellular components: nucleus (DNA), endoplasmic reticulum (ER), nucleoli, cytoplasmic RNA (RNA), actin, Golgi, plasma membrane (AGP) and mitochondria (Mito)<sup>1</sup>. Promisingly, pharmaceutical companies such as Recursion Pharmaceuticals have implemented Cell Painting to support clinical stage pipelines<sup>6</sup>.

Image-based assays suffer the technical limitation of having a finite number of imaging channels due to the requirement to avoid spectral overlap. Typically, there is a maximum of six stains which can be applied simultaneously across five imaging channels<sup>7</sup>, thus hindering the ability to capture an even greater quantity of morphological information from other unstained subcellular compartments.

Advances in deep learning and computer vision have rapidly advanced image-based profiling and is expected to accelerate drug discovery<sup>2</sup>. It is possible to use transmitted light images such as brightfield as an input to a convolutional neural network to generate the corresponding fluorescent images—so called *In Silico* labelling<sup>8</sup>. Brightfield imaging is ubiquitous in cell microscopy and has been proposed as an informative imaging modality in itself. While brightfield imaging overcomes many drawbacks of fluorescent labelling, it lacks the specificity and clear separation of the structures of interest. Deep learning methods have shown promise in augmenting the information available in brightfield when trained with fluorescent signals as a domain transfer problem<sup>9</sup>.

One way to achieve this is with conditional<sup>10</sup> generative adversarial networks (GANs)<sup>11</sup>, introduced by Isola in 2016 (also known as *pix2pix*)<sup>12</sup>. From the input samples of an underlying unknown joint distribution of multi-modal data, the generator network can be trained to generate complementing paired data. The advantage of the conditional GAN is to overcome the limitations of a pixel-wise loss function such as L1<sup>13</sup>, used in many standard U-Net models.

In this study we investigate the utility of label free Cell Painting. Image-to-image approaches in Cell Painting are rare, with only known instance of GANs specifically for Cell Painting using Nucleoli, cytoplasmic RNA, Golgi plasma membrane and F-actin as an input to predict the other two fluorescent channels<sup>14</sup>. Fluorescent channels

<sup>1</sup>Department for Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, UK. <sup>2</sup>Discovery Sciences, R&D, AstraZeneca, Cambridge Science Park, Milton Road, Cambridge CB4 0WG, UK. <sup>3</sup>Discovery Sciences, R&D, AstraZeneca, Alderley Park, Cheshire SK10 4TG, UK. <sup>4</sup>Discovery Sciences, R&D, AstraZeneca, Gothenburg, Pepparedsleden 1, 431 50 Mölndal, Sweden. <sup>5</sup>Downing College, Cambridge CB2 1DQ, UK. <sup>6</sup>These authors contributed equally: Riku Turkki and Yin Hai Wang. ✉email: jc856@cam.ac.uk

contain significantly more information than a brightfield z-stack, therefore predicting full fluorescent channels from brightfield images is a significantly more challenging task.

To the best of our knowledge this work is the first to demonstrate a label-free, five-channel Cell Painting replication from a transmitted light input modality such as brightfield. In this paper we have investigated the quality of the Cell Painting features extracted from the model-predicted image channels by correlating with the ground truth features. Additionally, we attempt to replicate clustering by treatment group on a test sample of 273 fields of view from two different batches. This was achieved by training two models on a large, multi-batch dataset: a U-Net trained with L1 loss, and the same U-Net trained as the generator of a conditional Wasserstein<sup>15</sup> GAN. In order to fully interrogate the utility of the label-free Cell Painting image channels, a comparison of both image-level metrics and morphological feature predictions between the two models is presented, contextualizing the non-normalized image metrics<sup>16</sup> and the feature-level evaluation results. We have aimed to provide a greater understanding to what extent convolutional neural network-based approaches may be able to assist or replace fluorescent staining in future studies.

## Materials and methods

**Cell culture and seeding.** A U-2 OS cell line was sourced from AstraZeneca's Global Cell Bank (ATCC Cat# HTB-96). Cells were maintained in continuous culture in McCoy's 5A media (#26600023 Fisher Scientific, Loughborough, UK) containing 10% (v/v) fetal bovine serum (Fisher Scientific, #10270106) at 37 °C, 5% (v/v) CO<sub>2</sub>, 95% humidity. At 80% confluency, cells were washed in PBS (Fisher Scientific, #10010056), detached from the flask using TrypLE Express (Fisher Scientific, #12604013) and resuspended in media. Cells were counted and a suspension prepared to achieve a seeding density of 1500 cells per well using a 40 µL dispense volume. Cell suspension was dispensed directly into assay-ready (compound-containing) CellCarrier-384 Ultra microplates (#6057300 Perkin Elmer, Waltham, MA, USA) using a Multidrop Combi (Fisher Scientific). Microplates were left at room temperature for 1 h before transferring to a microplate incubator at 37 °C, 5% (v/v) CO<sub>2</sub>, 95% humidity for a total incubation time of 48 h.

**Compound treatment.** All chemical compounds were sourced internally through the AstraZeneca Compound Management Group and prepared in stock solutions ranging 10–50 mM in 100% DMSO. Compounds are transferred into CellCarrier-384 Ultra microplates using a Labcyte Echo 555T from Echo-qualified source plates (#LP-0200 Labcyte, High Wycombe, UK). Compounds were tested at multiple concentrations; either 8 concentration points at threefold (half log) intervals or 4 concentration points at tenfold intervals (supplementary Table A contains concentration ranges for compound collections tested). Control wells situated on each plate consisted of neutral (0.1% DMSO v/v) and positive controls (63 nM mitoxantrone, a known and clinically-used topoisomerase inhibitor and DNA intercalator resulting in a cytotoxic phenotype). Compound addition to microplates was performed immediately prior to cell seeding to produce assay-ready plates.

**Cell staining.** The Cell Painting protocol was applied according to the original method<sup>1</sup> with minor adjustments to stain concentrations. Hank's balanced salt solution (HBSS) 10× was sourced from AstraZeneca's media preparation department and diluted in dH<sub>2</sub>O and filtered with a 0.22 µm filter. MitoTracker working stain was prepared in McCoy's 5A medium. The remaining stains were prepared in 1% (w/v) bovine serum albumin (BSA) in 1× HBSS (Table A supplementary).

Post incubation with compound, media was evacuated from assay plates using a Blue®Washer centrifugal plate washer (BlueCatBio, Neudrossenfeld, Germany). 30 µL of MitoTracker working solution was added and the plate incubated for a further 30 min at 37 °C, 5% CO<sub>2</sub>, 95% humidity. Cells were fixed by adding 11 µL of 12% (w/v) formaldehyde in PBS (to achieve final concentration of 3.2% v/v). Plates were incubated at room temperature for 20 min then washed using a Blue®Washer. 30 µL of 0.1% (v/v) Triton X-100 in HBSS (#T8787 Sigma Aldrich, St. Louis, MO, USA) solution was dispensed and incubated for a further 20 min at room temperature followed by an additional wash. 15 µL of mixed stain solution was dispensed, incubated for 30 min at room temperature then removed by washing. Plates were sealed and stored at 4 °C prior to imaging.

**Imaging.** Microplates were imaged on a CellVoyager CV8000 (Yokogawa, Tokyo, Japan) using a 20× water-immersion objective lens (NA 1.0). Excitation and emission wavelengths are as follows for fluorescent channels: DNA (ex: 405 nm, em: 445/45 nm), ER (ex: 488 nm, em: 525/50 nm), RNA (ex: 488 nm, em: 600/37 nm), AGP (ex: 561 nm, em: 600/37 nm) and Mito (ex: 640 nm, em: 676/29 nm). The three brightfield slices are from different focal z-planes; within, 4 µm above and 4 µm below the focal plane. Images were saved as 16-bit .tif files without binning (1996×1996 pixels).

**Pre-processing.** The selected images were bilinearly downsampled from 1996×1996 to 998×998 pixels to reduce computational overheads. Global intensity normalization was implemented to eliminate intensity differences between batches. For each channel, each 998×998 image was constrained to have a mean pixel value of zero and a standard deviation of one. Corrupted files or wells with missing fields were removed from the usable dataset and replaced with files from the corresponding batch.

**CellProfiler pipeline.** A CellProfiler<sup>17,18</sup> pipeline was utilized to extract image- and cell-level morphological features. The implementation followed the methodology of<sup>19</sup> and is chosen as a representative application of Cell Painting. The features extracted from this pipeline are included in our GitHub repository and in Supplementary Table B. CellProfiler was used to segment nuclei, cells and cytoplasm, then extract morphological features from

	No. of batches	No. of 384-well plates	No. of images	Controls & treatment distribution	Collection description (no. of compounds)
Total available	19	140	150,000 (4 fields of view per well)	2.1% positive controls; 6.4% negative controls; 91.5% treatments	Mixed collection containing a high chemical diversity set with a range of physicochemical properties (~ 7,000); a phenotypic set (~ 1,000) and a set of pan-assay interference compounds (~ 1,000)
Training/Validation set	17	124	3,000 (1 field of view used per well)	Randomly sampled from above	Randomly sampled from above
Test set	2	16	273 (1 field of view used per well)	26 positive controls; 77 negative controls; 170 treatments	Publicly available phenotypic set (exclusively in test set)

**Table 1.** The composition of the Training, Validation and Test sets. Choosing data from many batches and experiments should make our models learn more robust, common features, which gives us the best chance of successful prediction on unseen data.

each of the channels. Single cell measurements of fluorescence intensity, texture, granularity, density, location and various other features were calculated as feature vectors.

**Morphological profile generation.** Features were aggregated using the median value per image. For feature selection, we adopted the following approach:

- Drop missing features—features with > 5% NaN values, or zero values, across all images
- Drop blocklisted features<sup>20</sup> which have been recognized as noisy features or generally unreliable
- Drop features with greater than 90% Pearson correlation with other features
- Drop highly variable features (> 15 SD in DMSO controls)

This process reduced the number of useable features to 611, comparable to other studies. We used the ground truth data only in the feature reduction pipeline to avoid introducing a model bias to the selected features.

**Training and test set generation.** We sampled from 17 of the 19 batches to select wells for training (Table 1). The remaining two batches were used to select the test set and were excluded in the training process. Compounds from the test set were comprised from a set of known pharmacologically active molecules, with a known observable phenotypic activity. We randomly sampled 3000 wells for model training and hyperparameter tuning, with the constraint to force an overall equal number of wells per batch. We randomly selected one field of view from each of the four fields in the well, which was the image used in the training set. For model tuning we used 90/10 splits sampled randomly from the training set, before using the full training set to train the final models. The test set contained 273 images and was chosen by sampling randomly within each treatment group across the two remaining batches (treatment breakdown in Table 1).

**U-net with L1 loss network architecture.** Our first model is based on the original U-Net<sup>21</sup>, a convolutional neural network which has proven very effective in many imaging tasks. U-Net architectures have been used to solve inverse reconstruction problems in cellular<sup>22</sup>, medical<sup>23</sup> and general imaging problems<sup>24</sup>. For segmentation tasks, out of the box U-Net based architectures such as nnU-Net<sup>25</sup> have been proven to perform very well even compared to state-of-the-art models.

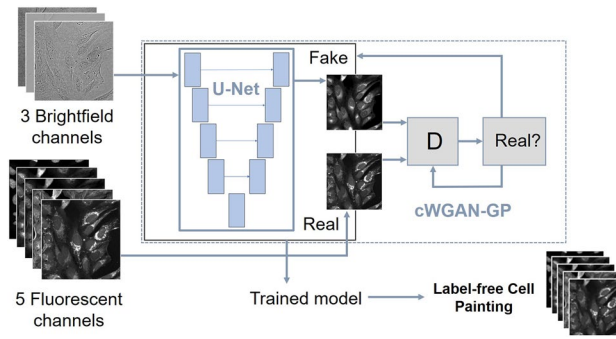
U-Nets involve a number of convolutions in a contracting path with down sampling or pooling layers, and an expansive path with up sampling and concatenations, allowing for retention of spatial information while learning detailed feature relationships. The network captures multi-scale features of images through different resolutions by going up and down the branches of the network.

We adapted the typical grayscale or RGB channel U-Net model to have 3 input channels and 5 output channels to accommodate our data. An overview of the model network and training is presented in Fig. 1. There were 6 convolutional blocks in the downsampling path, the first with 32 filters and the final with 1024 filters. Each block performed a 2d convolution with a kernel size of 3 and a stride of 1, followed by a ReLU then batch normalization operation. Between blocks a max pooling operation with a kernel size of 2 and a stride of 2 was applied for downsampling. The upsampling path was symmetric to the downsampling, with convolutions with a kernel of 2 and a stride of 2 applied for upsampling. The corresponding blocks in the contracting and expansive paths were concatenated as in the typical U-Net model. The final layer was a convolution with no activation or batch normalization. In total our network had  $31 \times 10^6$  trainable parameters.

For pairs of corresponding images  $\{(x_i, y_i)\}$ , where  $x_i$  is a 3-channel image from the brightfield input space and  $y_i$  is a 5-channel image from the real fluorescent output space, the loss function  $\mathcal{L}_{L1}$  for the U-Net model is:

$$\mathcal{L}_{L1} = \mathbb{E}_{x,y}[\|y - \hat{y}(x)\|_1] \quad (1)$$

where  $\hat{y}_i$  is the predicted output image from the network.



**Figure 1.** Summary of the models: U-Net trained on L1 loss and cWGAN-GP.

**Additional adversarial (GAN) training network architecture.** The second model we trained is a conditional<sup>10</sup> GAN<sup>11</sup>, where the generator network  $G$  is the same U-Net used in the first model. Typically, GANs have two components: the generator  $G$  and the discriminator  $D$ . The generator and the discriminator are simultaneously trained neural networks: the generator outputs an image and the discriminator learns to determine if the image is real or fake. For a conditional GAN, the objective function  $\mathcal{L}_{GAN}$  from a two-player minimax game is defined as:

$$\mathcal{L}_{GAN}(G, D) = \mathbb{E}_{x,y}[\log D(x, y)] + \mathbb{E}_x[\log(1 - D(x, G(x)))] \tag{2}$$

where  $G$  is trying to minimise this objective, and  $D$  is trying to maximize it:

$$\min_G \max_D \mathcal{L}_{GAN}(G, D) \tag{3}$$

Many GAN image reconstruction models follow the framework from Isola<sup>12</sup> where the model objective function  $G^*$  is, for example, a mix of both objective functions:

$$G^* = \arg \min_G \max_D \mathcal{L}_{GAN}(G, D) + \lambda \mathcal{L}_{L1}(G) \tag{4}$$

for some weighting parameter  $\lambda$ .

These constructions can be unstable and difficult to train, and this was the case for our dataset. To overcome difficulties with training we opted for a conditional Wasserstein GAN with gradient penalty<sup>26</sup> (cWGAN-GP) approach. This improved Wasserstein GAN was designed to stabilize training, useful for more complex architectures with many layers in the generator.

The Discriminator network  $D$ —alternatively the critic in the WGAN formulation—is a patch discriminator<sup>12</sup> with the concatenated brightfield and predicted Cell Painting channels as the eight-channel input. There were 64 filters in the final layer and there were three layers. For the convolutional operations the kernel size is 4 and the stride is 2. The output is the sum of the cross-entropy losses from each of the localized patches.

In Eq. (4), the L1 loss term enforces low-frequency structure, so using a discriminator which classifies smaller patches (and averages the responses) is helpful for capturing high frequency structures in the adversarial loss. For Wasserstein GANs the discriminator is called the critic as it is not trained to classify between real and fake, instead to learn a K-Lipschitz function to minimize the Wasserstein loss between the real data and the generator output.

Hence, for our conditional WGAN-GP-based construction we trained the generator to minimize the following objective:

$$\mathcal{L}_G(G, D) = \lambda_1 \mathcal{L}_{L1} - \lambda_e \mathbb{E}_{x,G(x)}[D(x, G(x))] \tag{5}$$

where  $\lambda_1$  is a weighting parameter for the L1 objective. We introduce  $\lambda_e$  as an adaptive weighting parameter to prevent the unbounded adversarial loss overwhelming L1:

$$\mathcal{L}_{L1} = \mathbb{E}_{x,y}[||y - G(x)||_1] \tag{6}$$

The critic objective, which the network is trained to maximize, is:

$$\mathcal{L}_D(G, D) = \mathbb{E}_{x,y}[D(x, y)] - \mathbb{E}_{x,G(x)}[D(x, G(x))] + \lambda_2 \mathcal{L}_{GP} \tag{7}$$

where  $\lambda_2$  is a weighting parameter for the gradient penalty term:

$$\mathcal{L}_{GP}(D) = \mathbb{E}_{x,\hat{x}}[(||\nabla_{\hat{x}} D(x, \hat{x})||_2 - 1)] \tag{8}$$

which is used to enforce the K-Lipschitz constraint.  $\hat{x}$  is from uniformly sampling along straight lines between pairs of points in the real data and generated data distributions<sup>27</sup>.

**Model training and computational details.** From the training set, random  $256 \times 256$  patches were cropped to serve as inputs to the network. No data augmentation was used. The models were trained on the AstraZeneca Scientific Computing Platform with a maximum allocation of 500G memory and four CPU cores. The PyTorch implementation of the networks and training is given in our GitHub repository. The U-Net model was trained with a batch size of 10 for 15,000 iterations (50 epochs). The optimizer was Adam, with a learning rate of  $2 \times 10^{-4}$  and weight decay of  $2 \times 10^{-4}$ . The total training time for the U-Net model was around 30 h.

The cWGAN-GP model was trained with a batch size of 4 for an additional 21,000 iterations (28 epochs). The generator optimizer was Adam, with a learning rate of  $2 \times 10^{-4}$  and  $\beta_1$  of 0 and  $\beta_2$  of 0.9. The critic optimizer was Adam with a learning rate of  $2 \times 10^{-4}$  and  $\beta_1$  of 0 and  $\beta_2$  of 0.9 and a weight decay of  $1 \times 10^{-3}$ . The generator was updated once for every 5 critic updates. The  $L1$  weight is  $\lambda_1 = 100$  and the gradient penalty weighting parameter is  $\lambda_2 = 10$ .  $\lambda_c = 1/\text{epoch}$ . The total training time for the cWGAN-GP model was an additional 35 h. For each model, the best performing epoch was selected by plotting the image evaluation metrics of the validation split for each epoch during training. Once the metrics stopped improving (or got worse), training was stopped as the model was considered to be overfit.

**Model inference.** Input and output images were processed in  $256 \times 256$  patches. The full-sized output images were stitched together with a stride of 128 pixels. This was chosen as half of the patch size so the edges met along the same line, which reduced the number of boundaries between tiles in the full-sized reconstructed image. Each pixel in the reconstructed prediction image for each fluorescent channel is the median value of the pixels in the four overlapping images. The full-sized, restitched images used for the metric and CellProfiler evaluations were  $998 \times 998$  pixels for each channel, examples of which are displayed in supplementary Figure B.

**Image-level evaluation metrics.** The predicted and target images were evaluated with five metrics: mean absolute error (MAE), mean squared error (MSE), structural similarity index measure (SSIM)<sup>28,29</sup>, peak signal-to-noise ratio (PSNR)<sup>29</sup> and the Pearson correlation coefficient (PCC)<sup>8,30</sup>. MAE and MSE capture pixel-wise differences between the images, and low values are favorable for image similarity. SSIM is a similarity measure between two images where for corresponding sub-windows of pixels, luminance, contrast, means, variances and covariances are evaluated. The mean of these calculations is taken to give the SSIM for the whole image. PSNR is a way of contextualizing and standardizing the MSE in terms of the pixel values of the image, with a higher PSNR corresponding to more similar images. PCC is used to measure the linear correlation of pixels in the images.

PSNR is normalized to the maximum potential pixel value, taken as 255 when the images are converted to 8-bit. For this dataset the PSNR appeared as high as the maximum 8-bit pixel values for each image was generally lower than the theoretical maximum value. Only SSIM can be interpreted as a fully-normalized metric, with values between 0 and 1 (1 being a perfect match).

**Morphological feature-level evaluation.** Pairwise Spearman correlations<sup>31</sup> between the features in test set data were calculated for each model, with the mean values for each feature group grouped into correlation matrices and visualized as heatmaps<sup>19</sup>. These features were split into several categories – area/shape, colocalization, granularity, intensity, neighbors, radial distribution and texture. We also visualized feature clustering using uniform manifold approximations (UMAPs)<sup>32</sup>, implemented in python using the UMAP package<sup>33</sup> with the parameters:  $n\_neighbors = 15$ ,  $min\_dist = 0.8$ ,  $n\_components = 2$ . Further details are in the GitHub repository.

**Toxicity prediction.** We normalized the morphological profiles to zero mean and standard deviation of one and classified the compounds with K-NN-classifier ( $k = 5$ , Euclidian distance) into two groups using the controls as training label (positive controls were used as an example of toxic phenotype). To account for the unbalance between the number of positive and negative controls, we sampled with a replacement equal number of profiles from both categories for training and ran the classifier 100 times and used majority voting for the final classification.

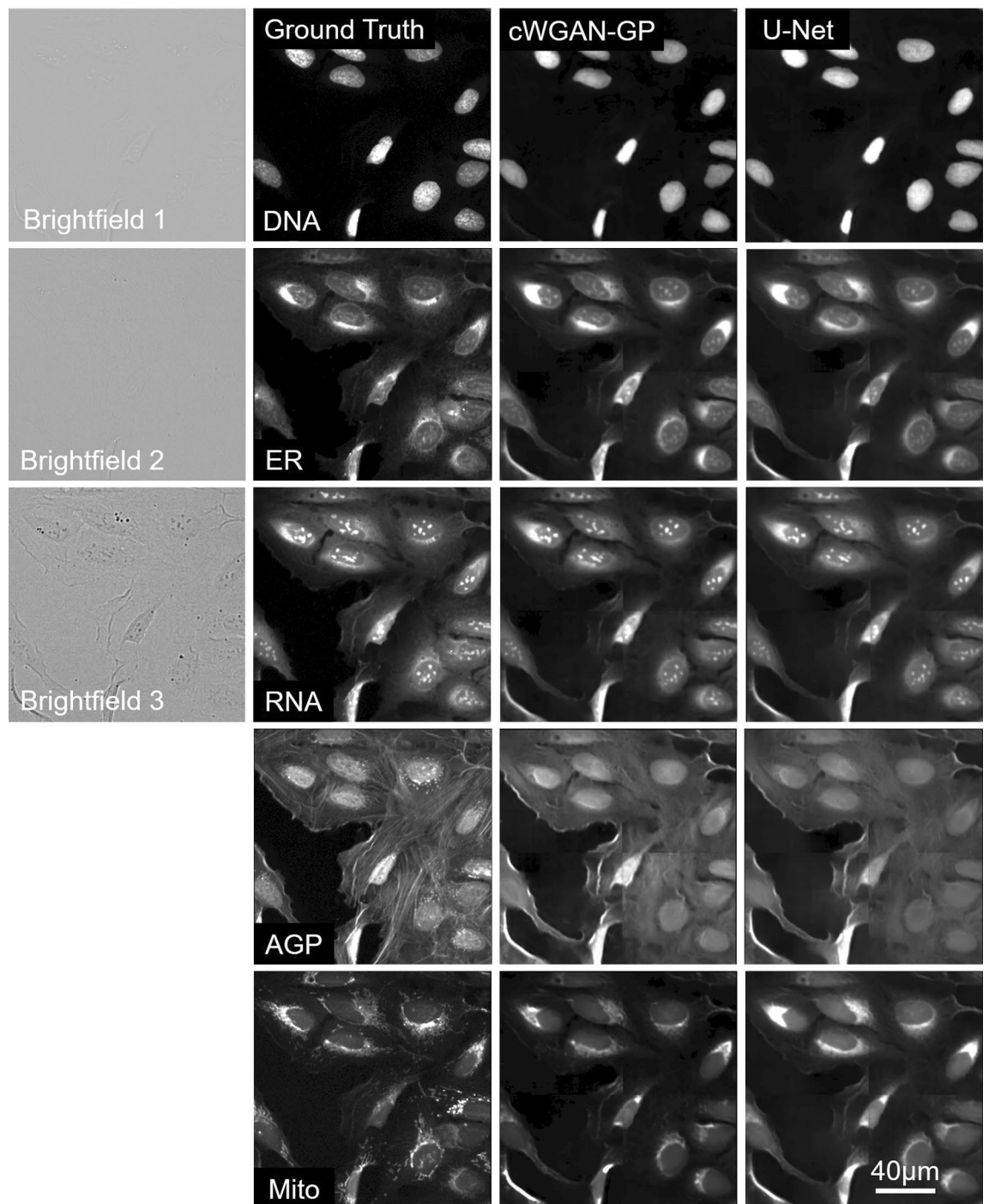
**Ethics approval/Consent to participate.** The authors declare this study does not require ethical approval.

## Results

To systematically investigate the utility of the label-free prediction of Cell Painting from brightfield images, we conducted the evaluation on three separate levels: image-level, morphological feature-level, as well as a downstream analysis on the profile-level to identify toxic compounds. We evaluated the two models on the test set of 148 unique compounds, 26 positive control wells (mitoxantrone), and 77 negative control wells (DMSO) that represented two experimental batches and 12 plates (Table 1). Examples and descriptions of cells in these different wells are presented in supplementary Figure A.

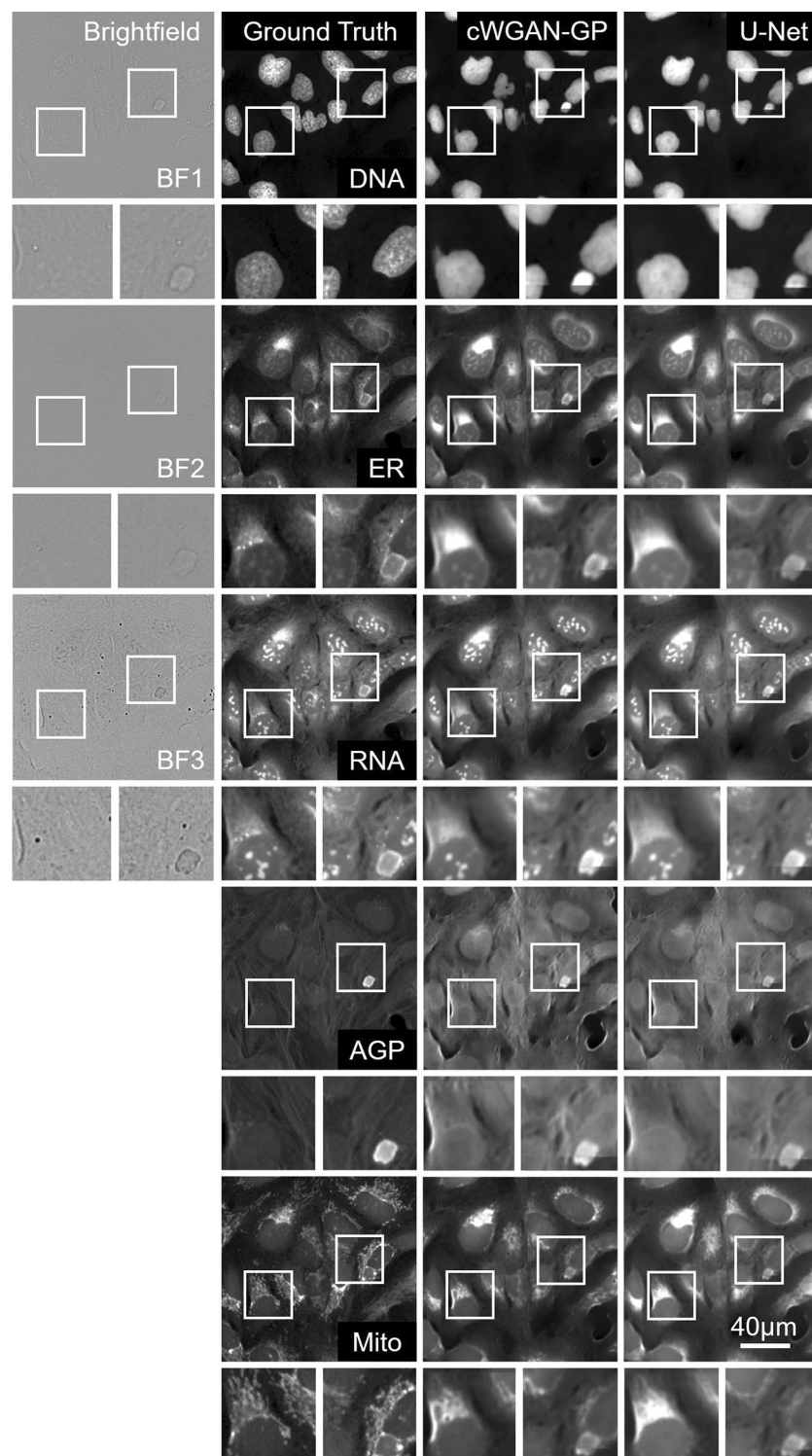
**Image-level evaluation.** The mean values of the image-level metrics for the U-Net/cWGAN-GP models respectively were SSIM: 0.42/**0.44**, PSNR: 51.7/**52.9**, MSE: 0.46/**0.45**, MAE: 0.41/**0.39**, and PCC: **0.84**/0.84, with the superior model in bold. In general, the cWGAN-GP generated images yield stronger image-level metrics, however the difference between the U-Net and cWGAN-GP metrics was small and not statistically significant (within one standard deviation). Visual inspection of the model generated images with the ground truth revealed strong resemblance (Figs. 2, 3, 4). Most notably, both models struggled in predicting the fine structures in AGP and Mito channels, this was also reflected as lower image-level performance for the channels. Interest-





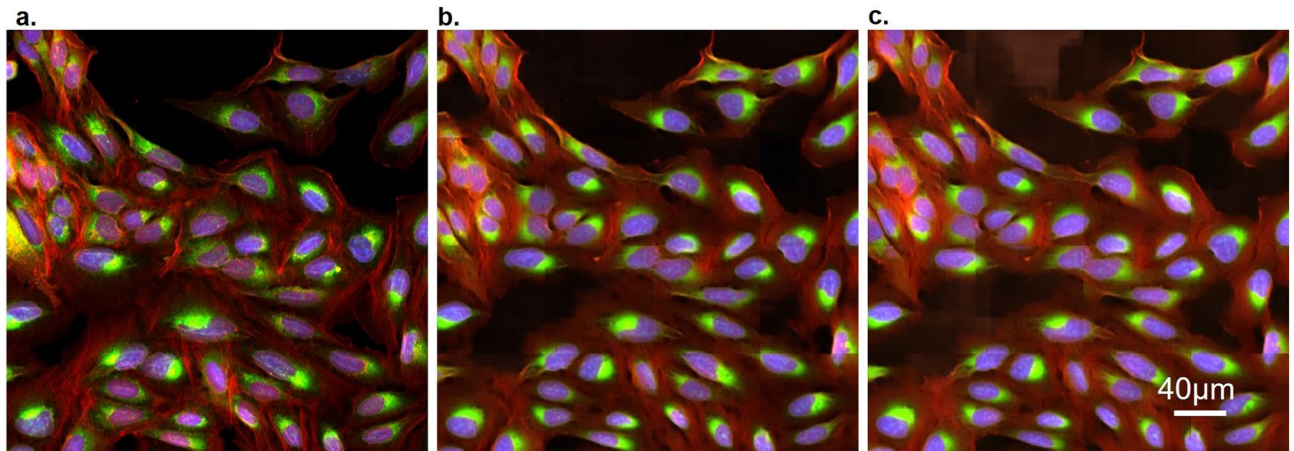
**Figure 2.** Cropped images of typical brightfield images, ground truth fluorescent, and predicted channels from the test dataset (Table 1) for the U-Net and cWGAN-GP models. Presented at scale of the input and output size used in the model networks ( $256 \times 256$  pixels), the first column contains the three brightfield z-planes which form the input, and the subsequent columns are the five fluorescent output channels. The stitching method has been applied to the model predicted images, and small artefacts of this process are visible. Images are independently contrasted for visualization.

ingly, only the plasma membrane of AGP channel was successfully predicted, whereas the models were not able to reproduce the actin filaments and the Golgi-apparatus structures. Similarly with the Mito channel, the models performed well overall in predicting the mitochondrial structures, but we observed lack in small granularity and detail that was present in the ground truth. In addition, we observed that the blocking effect from inference time sliding window was visible in the generated images. Nevertheless, DNA, ER, and RNA channels were better predicted, of which RNA achieved the best image-level performance.

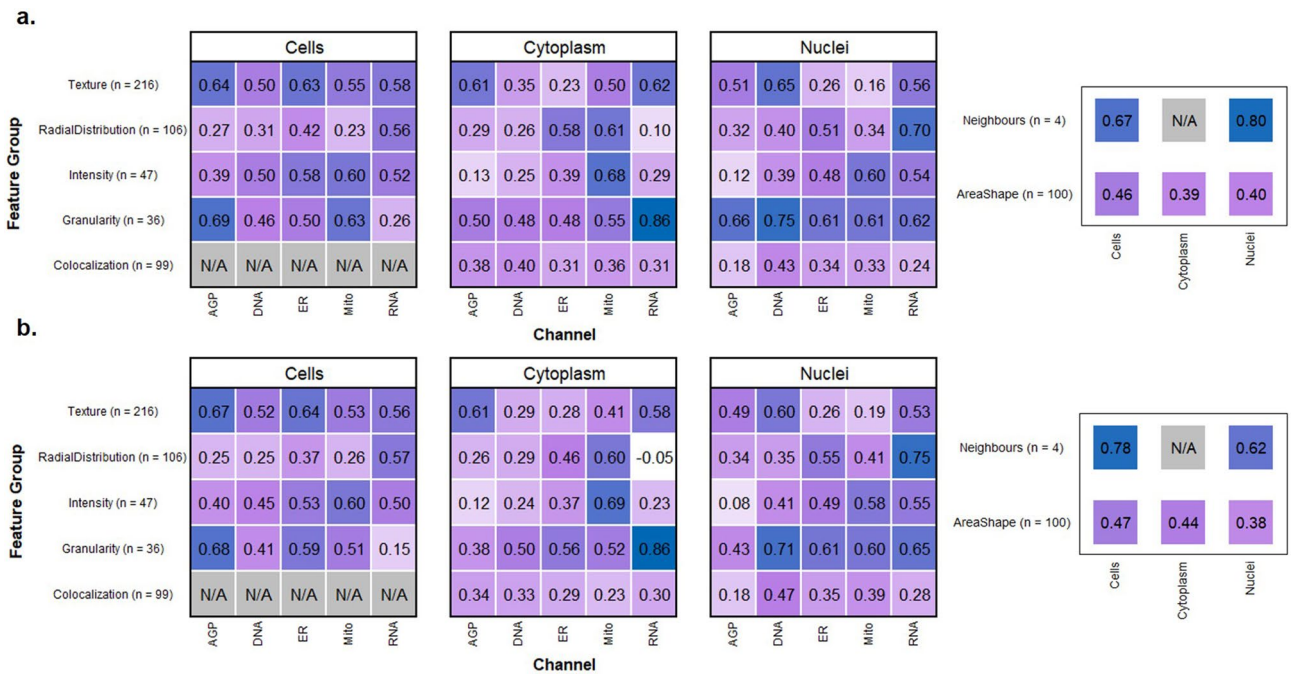


**Figure 3.** Two small insets for each image show a magnified view of the selected region of the cell and can be used to compare details between the ground truth and the two model outputs in each channel. Images are independently contrasted for visualization.

**Morphological feature-level evaluation.** Next, we extracted morphological features at the single-cell level and aggregated them to well-level profiles with the aim to further elucidate the utility of label-free Cell Painting. For consistency and to avoid any potential bias favoring the trained models, we chose to use the ground



**Figure 4.** Colored composite image of three channels from the test dataset (Table 1): AGP (red), ER (green) and DNA (blue). (a) Ground truth. (b) cWGAN-GP model output. (c) U-Net model output. Each image is 512 × 512 pixels.

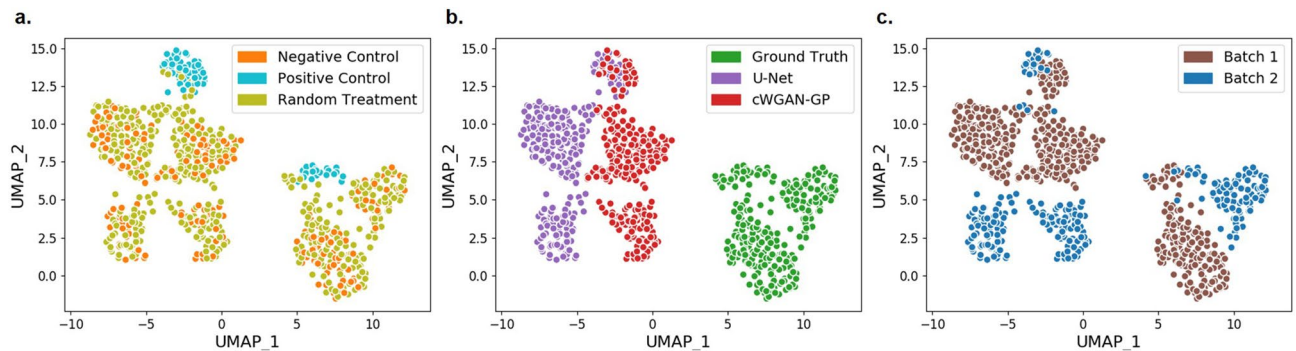


**Figure 5.** A systematic breakdown of correlation between morphological profiles. The features selected from the CellProfiler pipeline are presented as correlation heatmaps in (a) (U-Net) and (b) (cWGAN-GP). The correlations for each model are with the features extracted from the ground truth Cell Painting images and are aggregated by channel and feature group for cell, cytoplasm, and nuclei objects. The number of features for each feature group is also presented. U-Net mean = 0.43, cWGAN-GP mean = 0.45.

truth images for feature selection and reduced the profiles extracted from the predicted images accordingly. This resulted in 273 well-aggregated profiles consisting of 611 morphological features.

Using correlation analysis, we deconvoluted the profile similarities according to feature type, cell compartment, and imaging channels for both models (Fig. 5, Supplementary Figure D). Overall, many morphological features extracted from the generated images showed substantial correlation with those extracted from ground truth images. Examples of accurately reproduced (>0.6 correlation) feature groups across both models were texture measurements of the AGP channel in both cells and cytoplasm, intensity measurements of the Mito channel in the cytoplasm, and granularity and texture measurements of the DNA channel within the nuclei. The highest performing feature group were granularity measurements of the RNA channel within the cytoplasm (0.86 correlation in both models). Almost all features correlated positively with the ground truth, and only a small number of features showed close to zero or negative correlation, such as the radial distribution of RNA in cytoplasm and the intensity of the AGP channel in nuclei. The cell colocalization features were not calculated by CellProfiler however the cytoplasm and nuclei colocalization features represent the cell as a whole. A full





**Figure 6.** The UMAPs demonstrate that both models could reproduce the separation between treatments and batches seen in the ground truth features and that the clustering patterns are very similar. (a) shows each test image labelled by treatment type, (b) shows ground truth vs the two models and (c) highlights the difference between the two batches. cWGAN-GP was closer to the ground truth than U-Net, although the two models sit much closer to each other in feature space than either to the ground truth.

breakdown of the features extracted from each model is visualized in Supplementary Figures F and G, with the raw correlation values being provided in Supplementary Tables C and D.

The mean correlation of all the feature groups was 0.43 for the U-Net and 0.45 for cWGAN-GP, supporting the earlier finding of cWGAN-GP's superiority in the image-level evaluation. The feature breakdown by cell group was: 169 cell features (mean correlation: 0.48/0.50), 217 nuclei features (0.43/0.41), and 225 cytoplasm features (0.38/0.42). The feature breakdown by feature type was, in descending order of best to worst mean correlation: 4 neighbors (0.70/0.74), 36 granularity (0.53/0.55), 216 texture (0.47/0.49), 47 intensity (0.46/0.48), 103 area/shape (0.42/0.43), 106 radial distribution (0.38/0.39) and 99 colocalization (0.31/0.33). By channel, the mean feature correlations were: RNA: 0.47/0.49, AGP: 0.43/0.44, DNA: 0.41/0.44, Mito: 0.40/0.43, ER: 0.42/0.40.

The mean correlations for the top 10% of the selected features were 0.64/0.65. The number of features with a correlation greater than 0.8 were 26/30 (U-Net/cWGAN-GP), and for both models all feature groups and cell compartments had at least one feature with such a strong correlation, except for the colocalization feature group. For the cWGAN-GP model the breakdown of the 30 features with greater than 0.8 correlation was as follows: 12 cell, 11 nuclei, 7 cytoplasm, and this included the feature types: 11 texture, 7 radial distribution, 5 area/shape, 4 granularity, 2 texture, 1 neighbors, 0 colocalization.

Using uniform manifold approximation (UMAP), we visualized the high-dimensional morphological profiles for identification of underlying data structures (Fig. 6). UMAP was chosen as it is typically used for visualizing multidimensional morphological data and is regularly applied in the analysis of Cell Painting profiles<sup>2,19</sup>. We observed that all three image sources (ground truth, U-Net, and cWGAN-GP) were separated in the common feature space due to the difference in the ground truth and predicted images. Despite this separation, we also observed that all the image sources maintained the overall data structure. The clear batch effect visible in the ground truth is also evident in the predicted images, and similarly the clustering of positive and negative control wells respectively is retained, indicating successful model performance. Features extracted from the cWGAN-GP model lie closer to the ground truth features than the U-Net extracted features however features from both models are closer to each other than either are to the ground truth. We repeated the analysis using Principal Component Analysis (PCA)<sup>34</sup> and observed the same behavior (Supplementary Figure C).

**Profile-level evaluation.** Conclusions made from the label-free prediction of Cell Painting images should show agreement with ground truth datasets to be of experimental value. We therefore performed a series of analyses to identify compounds that elicit a comparable toxic phenotype to an established positive control compound, mitoxantrone. Our UMAP (Fig. 6) showed that, as would be expected, most of the compounds showed greater resemblance to negative (DMSO) controls. Promisingly, some compounds clustered with our positive control compound. To identify these compounds, we trained a K-NN classifier using the control profiles as our training set. The classification resulted in identification of eight compounds in the ground truth, five in the cWGAN-GP model, and eight in the U-Net model profiles (Supplementary Figure E). The U-Net model achieved a sensitivity of 62.5% and specificity of 98.0% in toxicity classification whilst the respective values for cWGAN-GP were 50.0% and 99.3%.

## Discussion

As the first full prediction of five-channel Cell Painting from brightfield input, we have presented evidence that label-free Cell Painting is a promising and practicable approach. We show our model can perform well in a typical downstream application, and that many label-free features from the predicted images correlate well with features from the stained images. In addition, we see success in clustering by treatment type from the features. We present indications of which channels and biomarkers can be satisfactorily by testing our model predictions with a comprehensive segmentation-based feature extraction profiling methodology.

Imaging channel	Model	SSIM	PSNR	MSE	MAE	PCC
DNA	U-Net	0.38 ± 0.13	51.8 ± 1.7	0.45 ± 0.12	0.41 ± 0.05	0.90 ± 0.06
	<b>cWGAN-GP</b>	<b>0.39 ± 0.14</b>	<b>52.4 ± 1.5</b>	<b>0.39 ± 0.10</b>	<b>0.39 ± 0.04</b>	<b>0.90 ± 0.06</b>
ER	U-Net	<b>0.49 ± 0.09</b>	<b>48.0 ± 1.3</b>	<b>1.07 ± 0.30</b>	<b>0.58 ± 0.09</b>	0.86 ± 0.06
	cWGAN-GP	0.48 ± 0.09	47.8 ± 1.3	1.12 ± 0.32	0.61 ± 0.09	<b>0.87 ± 0.06</b>
RNA	U-Net	0.59 ± 0.07	56.0 ± 1.6	0.18 ± 0.09	0.27 ± 0.06	0.92 ± 0.06
	<b>cWGAN-GP</b>	<b>0.60 ± 0.07</b>	<b>56.3 ± 1.7</b>	<b>0.17 ± 0.10</b>	<b>0.26 ± 0.06</b>	<b>0.92 ± 0.06</b>
AGP	U-Net	0.35 ± 0.08	52.4 ± 1.8	0.27 ± 0.10	0.39 ± 0.06	0.80 ± 0.07
	<b>cWGAN-GP</b>	<b>0.37 ± 0.09</b>	<b>54.7 ± 1.3</b>	<b>0.23 ± 0.09</b>	<b>0.34 ± 0.06</b>	<b>0.82 ± 0.06</b>
Mito	U-Net	0.31 ± 0.06	50.5 ± 1.1	<b>0.34 ± 0.08</b>	0.42 ± 0.05	<b>0.74 ± 0.08</b>
	<b>cWGAN-GP</b>	<b>0.35 ± 0.06</b>	<b>53.1 ± 1.4</b>	0.34 ± 0.12	<b>0.36 ± 0.04</b>	0.70 ± 0.08

**Table 2.** Image metrics for each channel for the two models. The best performing model for each channel metric is highlighted in bold.

Training with an adversarial scheme using a conditional GAN approach has been shown to enhance performance in virtual staining tasks<sup>35</sup>. In preliminary experiments, we tested different GAN models and adversarial weightings, but we chose cWGAN-GP due to its stability of training on our dataset. Increasing the adversarial component of the training objective function resulting in undesirable artifacts in the images which were not true representations of the cells. Our results demonstrate that incorporating adversarial loss results in a small increase in performance over L1 loss based on the pixel-wise evaluation metrics in all channels except ER. Even though the difference in metric values are mostly within one standard deviation, the values are calculated for images generated from the same set of brightfield images (for each model), so it is meaningful that cWGAN-GP achieves superior performance for 18 out of 25 metrics across the channels (Table 2). In domain transfer problems such as ours, the finer details in some of the predicted channels can be obstructed and blurred with a pixel-wise loss function such as L1 loss<sup>36</sup>. Just a small performance difference is expected as the network and loss function for training both models were very similar, and in the cWGAN-GP model the adversarial component of loss is weighted relatively low compared to L1.

In addition to the metric evaluation, features from images from the cWGAN-GP model show an increase in performance over the U-Net model, with slightly higher mean correlations to the ground truth (Fig. 5). The strong correlation of specific feature groups increases confidence that extracted morphological feature data from predicted images can be used to contribute overall morphological profiles of perturbations.

Biologically, it is expected that correlations of feature groups within the DNA channel are higher within the nuclei compartment than either cytoplasm or cells since the nuclei compartment is morphologically very distinct from the cytoplasmic region. The high correlation of radial distribution features in the RNA channel suggests that successful visualization of the nucleoli within the nuclear compartment has a large effect on this particular feature group. AGP and Mito channels both contain small and subtle cellular substructures which are typically less than two pixels wide. We postulate that this fine scale information is not present in the brightfield image in our data, making accurate image reproduction of the AGP and Mito channels very challenging regardless of model choice. There are known limitations of brightfield imaging which will always restrict a domain transfer model with brightfield input. For example, brightfield images can display heterogeneous intensity levels and poor contrast. Segmentation algorithms have been proven to perform poorly when compared with fluorescence channels, even after illumination correction methods have been applied<sup>27</sup>.

In UMAP feature space (Fig. 6), as well as in the PCA analysis (Supp Figure C), the features extracted from model-predicted images do not overlap with the ground truth features. This highlights limitations of the model but also the challenge of batch effect—it is not expected for the model to exactly predict an unseen batch. The relevant structure of the data is maintained although not absolute values. This is also notable at the image level, for example the high MSE is likely due to the batch effect causing a systematic difference in pixel values between the training and test batches. Similarity metrics such as SSIM may be more informative in this instance. It is notable that ground truth features from different batches also sit in non-overlapping feature space in our UMAPs. Within batches, the negative controls and treatments would form sub-clusters depending on the treatments in a larger dataset, however we acknowledge our test set was relatively small, resulting in minimal sub-clustering.

Other studies which have used U-Net based models to predict fluorescence from transmitted light or brightfield have evaluated their performance on pixel-level metrics such as PCC<sup>15</sup>, SSIM<sup>28,29</sup> and PSNR<sup>37</sup>. The mean PCC of all channels in our test set is 0.85 (using the best model for each channel), a value which compares favorably with prevailing work in fluorescent staining prediction<sup>8</sup>. In our data two channels (DNA, RNA) exceed a PCC of 0.90 for both models. However, absolute values of image metrics are heavily data dependent, so we present these metrics primarily for model comparison.

Such metrics are standard in image analysis but have some limitations for cellular data. Treating each pixel in the image equally is a significant limitation as pixels representing the cellular structures are clearly more important than background (void) pixels<sup>37</sup>. Some channels such as the DNA channel are more sparse than other channels such as AGP, and as such the number of pixels of interest vs background pixels is different. Extracting features from scientific pipelines provides a more biologically relevant and objective evaluation to give deep learning methods more credibility and a greater chance of being practically employed in this field<sup>38</sup>.

Despite the small sample size, the models performed well at predicting compounds with phenotypes similar to that of the positive control compound mitoxantrone, with specificities of 98.0% (U-Net) and 99.3%

(cWGAN-GP). The sensitivities were 62.5% and 50% respectively. Lower sensitivity values suggest the models are not identifying some potentially toxic or active profiles. There are multiple morphological features present in ground truth images which are linked to toxicity and it is likely that the models cannot capture all of these accurately. The manifestation of cytotoxicity can be most prevalent in the Mito/AGP channels due to mitochondrial dysfunction and gross changes to cytoskeletal processes and structure. We report that the Mito and AGP channels were the least well predicted channels on similarity and correlation metrics (Table 2), thus could serve as a reason for loss of sensitivity. Future studies could focus on predicting these two channels in particular. It is also relevant to consider that the concentration of mitoxantrone used was deliberately chosen to elicit a milder cytotoxic phenotype with an enlarged cellular area (suggesting a cytostatic effect is occurring) rather than complete cell death and an absence of cells entirely. Very high specificities indicate that label-free Cell Painting is not introducing prediction errors which would lead to false positive identification of a mitoxantrone-like phenotype. This is expected for models with significantly more weight on L1 loss, rather than with a high adversarial weighting which could introduce artifacts or phantom structures into the predicted images.

Despite a lack of sensitivity, we have demonstrated that morphological features extracted from both models are capable of recapitulating a significant portion of the morphological feature space to result in a positive clustering to the chosen control compound, mitoxantrone. We disclose two compounds which are already in the public domain; glipizide, a clinically-used sulfonylurea compound for the treatment of type 2 diabetes, and GW-842470, a phosphodiesterase 4 inhibitor previously evaluated in clinical studies for the treatment of atopic dermatitis (discontinued).

Although comparative or superior results for phenotype classification could likely be achieved with an image classifier trained on the brightfield, as a label-free and image-to-image approach there is a lot of promise as a generalist model to perform multiple tasks. The simple visualization of Cell Painting channels is one approach to improve the interpretation of brightfield images. Our results highlight the rich information captured in the brightfield modality, which may currently be under-utilized in morphological profiling.

Our approach by itself cannot replace in full the fluorescent staining process and the information it contains, but we have provided evidence that it may be possible to replicate the information of some Cell Painting channels and feature groups, and that the brightfield modality by itself may be sufficient for certain experimental applications. Importantly, employing such methods may reduce time, experimental cost and enable the utilization of specific imaging channels for experiment-specific, non-generic fluorescent stains. We acknowledge that particular feature groups which predict poorly in our models (such as colocalization features) may result in an inability (or reduced sensitivity) to identify cellular phenotypes which are characterized solely by these features. In such experimental applications, the replacement of generic stains for phenotype- or target-relevant biomarkers may offer an effective solution to the standard Cell Painting protocol.

One limitation of this study is the dataset used. Future studies will require larger datasets with greater diversity in terms of compound treatments and collection sources. Matching the numbers of fields in our training and test sets to the size of a typical dataset used in drug discovery would allow for greater insight into the capabilities of label-free Cell Painting. International collaborations such as the Joint Undertaking of Morphological Profiling using Cell Painting (JUMP-CP) aim to further develop Cell Painting and provide a highly valuable public dataset for this use<sup>39</sup>. It is notable that we have evaluated predictions with downstream features the networks have not seen—simply extracted classically from the images. In future studies, incorporating feature information in the training of the network itself may drastically improve performance, and this task may be appropriate for a transfer learning approach<sup>40</sup>.

We also acknowledge that the brightfield imaging mode itself may restrict the quantity of information being input into our models. The rationale behind imaging at multiple focal planes in brightfield configuration was to visualize as much cellular substructure as possible. We frequently observed that not all cellular features are visible in a single focal plane, therefore taking information from a z-stack image set will increase the input content to the models. Alternative brightfield imaging modalities such as phase contrast microscopy or differential interference contrast (DIC) microscopy have previously been used for fluorescence channel prediction tasks<sup>22,41</sup> and can capture a wealth of cellular morphology. It was beyond the scope of this study to investigate the impact of using different brightfield imaging modalities on predictive model performance, yet it should be recognized that all brightfield approaches lack the ability to fully visualize small-scale cellular substructures to an extent<sup>14</sup>. We propose that brightfield imaging in a z-stack configuration serves as a typical process that is widely adopted across both academic and industry laboratories, therefore the methods presented herein are applicable to a range of instruments with varied imaging capabilities.

In summary, we propose a deep learning approach which can digitally predict Cell Painting channels without the application of fluorescent stains, using only the brightfield imaging modality as an input. Building upon previous work<sup>8,22</sup> we have predicted the five fluorescent channels and used these images to extract the associated groups of morphological features from a standard image analysis pipeline. We then used the features from the predicted images to assess how information-rich such images are. Finally, we have provided a critical evaluation of the predictions using morphological features extracted from images using CellProfiler analysis and the resulting compound profiles.

## Data availability

The datasets generated and analyzed in this study are not publicly available due to AstraZeneca licenses but are available from the corresponding author on reasonable request. Source code from this project is available on GitHub: <https://github.com/crosszmirski/Label-free-prediction-of-Cell-Painting-from-brightfield-images>.

Received: 5 November 2021; Accepted: 18 May 2022

Published online: 15 June 2022

## References

- 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).
- Chandrasekaran, S. N., Ceulemans, H., Boyd, J. D. & Carpenter, A. E. Image-based profiling for drug discovery: Due for a machine-learning upgrade?. *Nat. Rev. Drug Discovery* **20**, 145–159 (2021).
- Willis, C., Nyffeler, J. & Harrill, J. Phenotypic profiling of reference chemicals across biologically diverse cell types using the cell painting assay. *SLAS Discovery: Adv. Life Sci. R & D* **25**(7), 755–769 (2020).
- Pahl, A. & Sievers, S. The cell painting assay as a screening tool for the discovery of bioactivities in new chemical matter. *Methods Mol. Biol.* **1888**, 115–126 (2019).
- Rohban, M. H. *et al.* Systematic morphological profiling of human gene and allele function via Cell Painting. *Elife* **6**, e24060 (2017).
- Mullard, A. Machine learning brings cell imaging promises into focus. *Nat. Rev. Drug Discovery* **18**(9), 653–655 (2019).
- Haraguchi, T., Shimi, T., Koujin, T., Hashiguchi, N. & Hiraoka, Y. Spectral imaging fluorescence microscopy. *Genes Cells: Devot Mol Cell Mech* **7**(9), 881–887 (2002).
- Christiansen, E. M. *et al.* In silico labeling: Predicting fluorescent labels in unlabeled images. *Cell* **173**(3), 792–803.e19 (2018).
- Imboden, S. *et al.* Investigating heterogeneities of live mesenchymal stromal cells using AI-based label-free imaging. *Sci Rep* **11**, 6728 (2021).
- Mirza, M., Osindero, S. Conditional Generative Adversarial Nets. [arXiv:1411.1784](https://arxiv.org/abs/1411.1784) (2014)
- Goodfellow, I. *et al.* Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* **27**, 2672–2680 (2014).
- Isola, P., Zhu, J.Y., Zhou, T. & Efros, A.A. Image-to-image translation with conditional adversarial networks. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1125–1134 (2017)
- Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A. & Acosta, A., *et al.* Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 105–114 (2017)
- Nguyen, T. *et al.* Virtual organelle self-coding for fluorescence imaging via adversarial learning. *J. Biomed. Opt.* **25**(9), 096009 (2020).
- Arjovsky, M., Chintala, S. & Bottou, L. Wasserstein Generative Adversarial Networks. *International Conference on Machine Learning, ICML*, (2017).
- Sara, U., Akter, M. & Uddin, M. Image Quality Assessment through FSIM, SSIM, MSE and PSNR—A Comparative Study. *J. Comput. Commun.* **7**, 8–18 (2019).
- Carpenter, A. E. *et al.* Cell profiler: Image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**, R100 (2006).
- McQuin, C. *et al.* Cell Profiler 3.0: Next-generation image processing for biology. *PLoS Biol.* **16**, e2005970 (2018).
- Way, G. P. *et al.* Predicting cell health phenotypes using image-based morphology profiling. *Mol. Biol. Cell* **32**(9), 995–1005 (2021).
- Way, G.P. Blocklist Features - Cell Profiler. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.10255811.v3> (2019)
- Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. MICCAI 2015. Lecture Notes in Computer Science, vol 9351, Springer, Cham (2015)
- Ounkomol, C., Seshamani, S., Maleckar, M. M., Collman, F. & Johnson, G. R. Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy. *Nat. Methods* **15**(11), 917–920 (2018).
- Hyun, C. M., Kim, H. P., Lee, S. M., Lee, S. & Seo, J. K. Deep learning for undersampled MRI reconstruction. *Phys. Med. Biol.* **63**(13), 135007 (2018).
- Jin, K. H., McCann, M. T., Froustey, E. & Unser, M. Deep convolutional neural network for inverse problems in imaging. *IEEE Trans. Image Process.: Publ. IEEE Signal Process. Soc.* **26**(9), 4509–4522 (2017).
- Isensee, F., Jaeger, P. F., Kohl, S., Petersen, J. & Maier-Hein, K. H. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **18**(2), 203–211 (2021).
- Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V. & Courville, A.C. Improved training of Wasserstein GANs. In *Advances in Neural Information Processing Systems (NIPS)*, 5769–5779 (2017)
- Buggenthin, F. *et al.* An automatic method for robust and fast cell detection in bright field images from high-throughput microscopy. *BMC Bioinform.* **14**, 297 (2013).
- Wang, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**(4), 600–612 (2004).
- Horé, A. & Ziou, D. Image Quality Metrics: PSNR vs. SSIM, *2010 20th International Conference on Pattern Recognition*, 2366–2369 (2010)
- Schober, P., Boer, C. & Schwarte, L. A. Correlation coefficients: appropriate use and interpretation. *Anesth. Analg.* **126**(5), 1763–1768 (2018).
- Mukaka, M. M. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med. J.* **24**(3), 69–71 (2012).
- McInnes, L., Healy, J., Saul, N. & Großberger, L. UMAP: Uniform manifold approximation and projection. *J. Open Source Softw.* **3**, 861 (2018).
- UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction, [Internet], Copyright 2018, Leland McInnes Revision 23b789e0, [cited October 2021]. Available from <https://umap-learn.readthedocs.io/en/latest/#>
- Jolliffe, I.T. Mathematical and statistical properties of population principal components. *Princ. Compon. Anal.* 10–28 (2002)
- Wieslander, H., Gupta, A., Bergman, E., Hallström, E. & Harrison, P. J. Learning to see colours: Biologically relevant virtual staining for adipocyte cell images. *PLoS ONE* **16**(10), e0258546 (2021).
- Ghodrati, V. *et al.* MR image reconstruction using deep learning: Evaluation of network structure and loss functions. *Quant. Imaging Med. Surg.* **9**(9), 1516–1527 (2019).
- Pambrun, J.F. & Noumeir, R. Limitations of the SSIM quality metric in the context of diagnostic imaging. *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2960–2963 (2015)
- Cheng, S., Fu, S., Kim, Y.M., Song, W., Li, Y., Xue, Y., *et al.* Single-cell cytometry via multiplexed fluorescence prediction by label-free reflectance microscopy. *Sci. Adv.* **7**(3), eabe0431 (2021)
- JUMP-Cell Painting Consortium, Joint Undertaking in Morphological Profiling, [Internet], 2021 Broad Institute, [cited October 2021]. Available from <https://jump-cellpainting.broadinstitute.org/>
- Bengio, Y. Deep Learning of Representations for Unsupervised and Transfer Learning. *ICML Unsupervised and Transfer Learning* (2012)
- Liu, Y., Yuan, H., Wang, Z. & Ji, S. Global pixel transformers for virtual staining of microscopy images. *IEEE Trans. Med. Imaging* **39**(6), 2256–2266 (2020).



## Acknowledgements

The authors would like to thank the AZ UKCCB team for the provision of the cell line. We also thank Charles-Hugues Lardeau, Diane Smith and Hannah Semple for their assistance in generating the imaging dataset used in this publication.

## Author contributions

Y.W., R.T., J.C.-Z. and C.-B.S. performed study concept and design; J.C.-Z. and R.T. performed the computational experiments, methodology development, data interpretation and statistical analysis; E.M. and G.W. conducted laboratory work and imaging; All the authors contributed to the writing, review and revision of the paper. All the authors read and approved the final paper.

## Funding

J.C.-Z. is funded by the UKRI-BBSRC DTP (UK Research and Innovation and the Biotechnology and Biological Science Research Council Doctoral Training Partnership) studentship grant 21-3688 and AstraZeneca. C.-B.S. acknowledges support from the Philip Leverhulme Prize, the Royal Society Wolfson Fellowship, the EPSRC grants EP/S026045/1 and EP/T003553/1, EP/N014588/1, EP/T017961/1, the Wellcome Trust 215733/Z/19/Z and 221633/Z/20/Z, the Leverhulme Trust project Unveiling the invisible, the European Union Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 777826 NoMADS, the Cantab Capital Institute for the Mathematics of Information and the Alan Turing Institute. Authors E.M., G.W., R.T., Y.W. are employees of AstraZeneca. AstraZeneca provided the funding for this research and provided support in the form of salaries for the authors, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the 'Author Contribution Statement' section.

## Competing interests

J.C.-Z., E.M., G.W., R.T. and Y.W. are employees of AstraZeneca plc. The remaining authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-12914-x>.

**Correspondence** and requests for materials should be addressed to J.O.C.-Z.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022