Artificial Intelligence in the Life Sciences 10 (2026) 100181

Contents lists available at ScienceDirect

Artificial Intelligence in the Life Sciences

journal homepage: www.elsevier.com/locate/ailsci

Research article

Single-cell morphological profiling reveals insights into programmed cell death

Benjamin M. Frey [a,*], David Holmberg [a], Petter Byström [b], Ebba Bergman [a], Polina Georgiev [a], Martin M. Johansson [a], Patrick Hennig [a], Jonne Rietdijk [a], Dan Rosén [a], Jordi Carreras-Puigvert [a,b], Ola Spjuth [a,b]

[a] Department of Pharmaceutical Biosciences and Science for Life Laboratory, Uppsala University, Uppsala, Sweden
[b] Pixl Bio AB, Uppsala, Sweden

ARTICLE INFO

Dataset link: https://github.com/pharmbio/sc_celldeath, https://doi.org/10.17044/scilifelab.28202864

Keywords:
Computational biology
Deep learning
machine learning
cell painting
single-cell

ABSTRACT

Analysis at the single-cell level is a powerful approach to study biological processes and responses to perturbations. However, its application in morphological profiling with phenomics remains underexplored. Here, we use the Cell Painting assay to investigate morphological effects of 53 small molecule compounds, associated with six distinct cell death mechanisms, across six concentrations in MCF7 cells. To compare single-cell and aggregated analysis strategies, we conduct both supervised and unsupervised evaluations aimed at identifying features linked to programmed cell death. We apply an energy distance as a metric to quantify morphological perturbation strength, enabling efficient filtering. Among three tested feature extraction methods, self-supervised DINO embeddings applied to single-cell data captured high-resolution morphological patterns. Focused analyses of apoptosis-inducing compounds revealed biological heterogeneity attributable to specific molecular targets and concentration-dependent effects, which were not apparent in aggregated profiles. In contrast, multi-class classification models for the six programmed cell death mechanisms trained on single-cell features achieved F1 scores of 79.86%, while models trained on aggregated features reached F1 scores of up to 89.97%. Our results highlight the advantages of single-cell data for unsupervised exploration and show that aggregated representations yield more robust and accurate performance in supervised models.

1. Introduction

Single-cell data has become a cornerstone of modern biological research, providing unprecedented insights into cellular heterogeneity, immune responses, and disease mechanisms. Specifically in the field of transcriptomics, single-cell analysis has enabled the identification of subpopulations within heterogeneous samples, aided in uncovering the effects of perturbations, and advancing our understanding of tumor biology [1–5]. Single-cell sequencing technologies have proven invaluable for exploring immune responses and elucidating disease mechanisms, particularly in the context of cancer biology [6]. Despite the growing impact of single-cell sequencing, the application of high-throughput single-cell data for characterizing specific programmed cell death (PCD) subtypes remains an open challenge.

The analysis pipeline for Cell Painting data begins with a feature extraction step that leverages tools like CellProfiler [7], a dedicated software that extracts pre-defined features from cellular images. CellProfiler has been instrumental in advancing morphological profiling by enabling researchers to extract detailed cellular characteristics [8–10]. However, the reliance on hand-crafted features has limitations, particularly as data scales and biological complexity increases. Recent advances in computer vision have introduced alternative approaches that leverage deep neural networks and optimized feature sets for automated feature extraction [11–13]. These approaches include convolutional neural networks (CNN) [14], weakly-supervised [15], and self-supervised models [12,16,17].

Self-supervised vision models, in particular, have shown great promise in morphological profiling as they do not require extensive data labeling. One example, DINO, employs two parallel networks to learn representations from different crops of the same image, enabling robust feature extraction. DINO has been successfully applied to Cell Painting data, outperforming existing methods in MOA prediction and toxicity estimation tasks [13,15,18–20]. Similarly, Masked Autoencoders, which use pixel masking and encoder–decoder architectures, have demonstrated state-of-the-art performance on Cell Painting

* Corresponding author.
E-mail address: benjamin.frey@helmholtz-munich.de (B.M. Frey).

https://doi.org/10.1016/j.ailsci.2026.100181
Received 16 June 2026; Received in revised form 6 August 2026; Accepted 17 August 2026
Available online 23 September 2026
2667-3185/© 2026 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

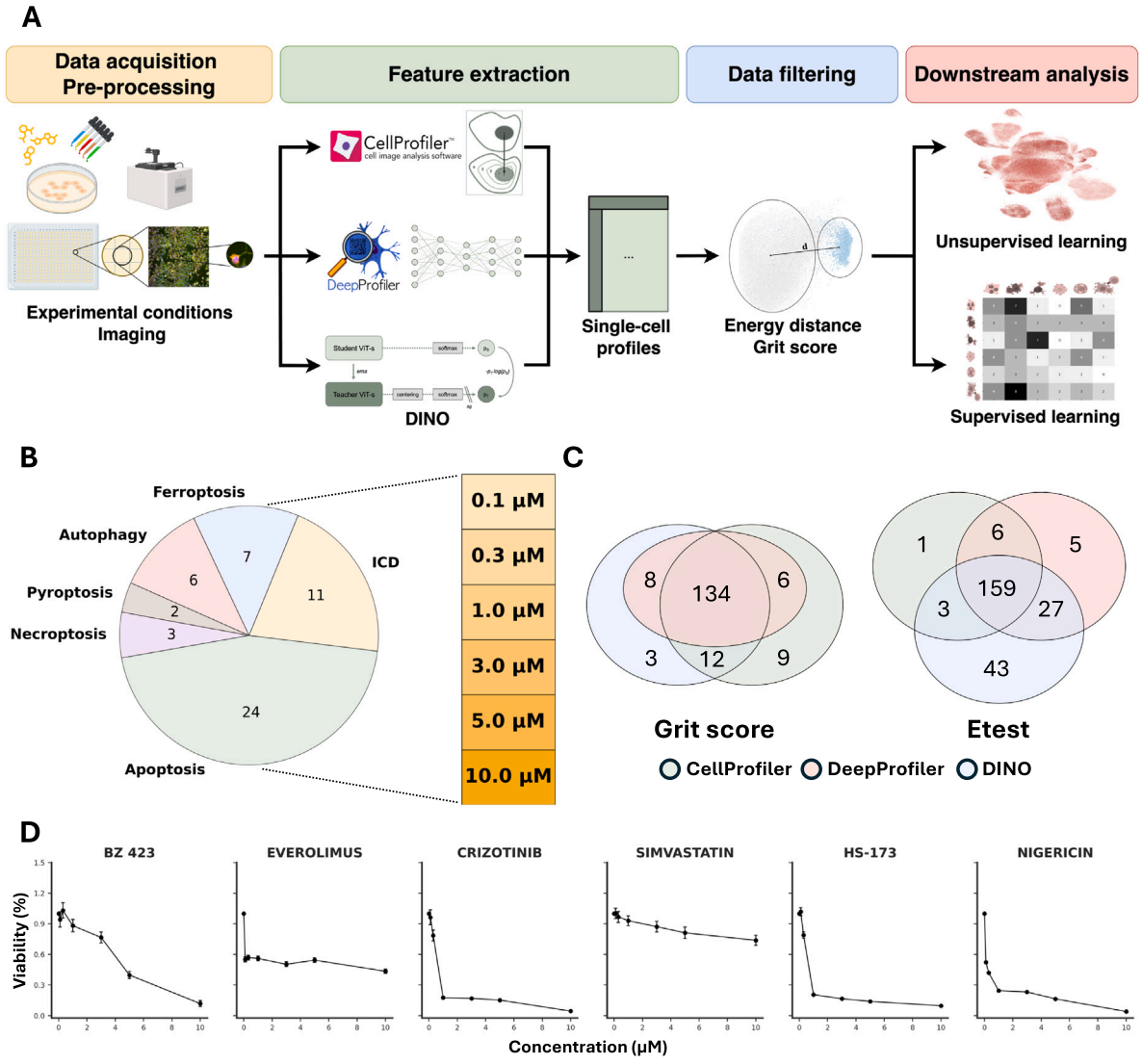

Fig. 1. Overview of experimental and analysis pipeline and filtering outcomes. (A) Schematic of the workflow used in this work. Experimental conditions are established, followed by imaging and feature extraction using CellProfiler, DeepProfiler, or DINO. Each extracted feature set undergoes separate data filtering and downstream analysis, including metric calculations, unsupervised and supervised learning. (B) Distribution of compound counts across different of cell death subtypes. Bar chart indicates concentrations used for the different compounds. The same concentrations were used in each compound. (C) Venn diagram depicting number of significant compound concentration sets between CellProfiler, DeepProfiler and DINO. Left diagram shows filtered grit sets, right diagram shows significant etest sets. (D) Dose-dependent viability per PCD subtype (1 randomly selected compound per subtype): Apoptosis, Autophagy, ICD, Ferroptosis, Necroptosis, Pyroptosis (from left to right). Mean ± standard error (n = 4 technical replicates).

datasets. Notably, Recursion Pharmaceuticals recently released an open version of a Masked Autoencoder model trained on millions of Cell Painting images, a first foundation model for Cell Painting data [12].

Despite these advancements, the majority of morphological profiling workflows where cells are perturbed using chemical or genetic perturbations aggregate features extracted from single-cell imaging into average profiles. By combining profiles from the same treatments and using mean/median cells as the main objects, outliers and data complexity can be reduced [15,21]. While this approach improves data separation at the treatment level, it effectively sacrifices information about cellular heterogeneity and subpopulations at the single-cell level. This trade-off limits the deeper biological characterization of individual cells, where heterogeneity can reveal essential insights into responses to perturbations.

Applications where Cell Painting experiments have utilized single-cell data in the analysis remain relatively limited. Early efforts focused on tasks like protein co-localization analysis, classification, and the characterization of blood and immune cell types [15,18–20,22]. One specific application where single-cell analysis from Cell Painting has not been explored extensively is PCD. Previous works from Lippincott et. al and Chen et. al have applied Cell Painting and high-content imaging on single-cells to study pyroptosis and the related inflammasome respectively [23,24]. In another work, Schorpp et al. classified different cell death subtypes using a CNN and larger image crops from four-channel Cell Painting experiments [25], covering two cell death subtypes (apoptosis, Ferroptosis, as well as non-treated cells) at seven inducing compounds per subtypes in five concentrations per compound. For their classification, they compared an end-to-end CNN with 512 x 512 crops as inputs with traditional machine learning models trained on 245-dimensional aggregated feature vectors.

In this manuscript we present a comprehensive exploration of six programmed cell death subtypes induced by 53 compounds in MCF-7 breast cancer cells, each tested at six different concentrations using five-channel Cell Painting. In order to assess the relevance of single-cell

morphological profiles for studying PCD subtypes, we apply unsupervised and supervised methods and compare the performance of three feature extraction approaches: CellProfiler [26], DeepProfiler [15], and the self-supervised transformer-based DINO (Fig. 1 A)

2. Methods and materials

2.1. Compounds and programmed cell death subtypes

53 compounds were selected where annotations spanned six distinct cell-death-associated categories: apoptosis, Necroptosis, Ferroptosis, Pyroptosis, as well as Autophagy and Immunogenic Cell Death (ICD) [24]. We note that these six categories are not equally validated as bona fide death mechanisms: apoptosis, necroptosis, ferroptosis, and pyroptosis are established forms of regulated cell death [27], whereas autophagy is a catabolic, context-dependent process that can be pro-survival rather than lethal, and immunogenic cell death (ICD) is defined by its downstream immunological consequences (DAMP exposure, adaptive immune priming) rather than by a distinct death morphology. As this study uses a monoculture MCF7 system without immune co-culture or DAMP assays, the autophagy and ICD labels reflect literature-established compound-mechanism annotations rather than confirmed death outcomes, and the six categories should be interpreted throughout as cell-death-associated compound annotations rather than validated death phenotypes. The compounds and PCD subtype annotations are shown in Fig. 1B; and with their associated targets in Supplemental Table S1.

Apoptosis, as the most extensively studied form of cell death, is characterized via cell shrinkage followed by fragmentation and development of apoptotic bodies [28]. Necroptosis describes cell death defined by membrane rupture and subsequent release of cellular contents. Necroptotic processes are induced by tumor necrosis factor (TNF) leading to a RIPK and caspase 8 activation [29]. Ferroptosis is driven by metabolic processes, specifically involving fatty acyl moieties dependent on iron [30]. It has been shown to be connected to different types of degenerative diseases involving the brain, heart and kidney. Pyroptosis presents a natural cellular defense mechanism induced by pathogen-associated molecular patterns (PAMPs) and damage-associated molecular patterns (DAMP) which trigger cell swelling, lysis and the release of pro-inflammatory cytokines, and shown to have cancer-related indications [31]. Autophagy is a process driven by the lysosomal machinery, resulting in the self-degradation of cellular components [32]. Autophagy is not intrinsically a death mechanism and can instead promote cell survival under stress; its inclusion here reflects recognized mechanistic links to regulated cell death pathways [24] rather than a validated lethal phenotype in this dataset. Immunogenic cell death (ICD), on the other hand, is a form of regulated cell death that elicits an adaptive immune response against antigens derived from dying neighboring cells, induced for instance by DNA damaging agents [33]. ICD is characterized by the release and surface exposure of specific DAMPs, acting as endogenous adjuvants to recruit highly specialized immune cells, such as CD8+ *T* cells, leading to the targeted destruction of the affected cell [34]. We did not directly assay DAMP release, calreticulin exposure, or immune cell activation in this study; the ICD annotation is based on established compound-mechanism literature rather than immunological validation within the present dataset.

2.2. Cell culture

MCF7 human breast cancer cells (Sigma-Aldrich #86012803) were cultured to passage +24 and confirmed to be Mycoplasma-free using the MycoAlert kit (Lonza #LT07-218). Cells were maintained at 37 °C, 5% CO_2, and 95% humidity in Dulbecco's Modified Eagle Medium (DMEM; Gibco #31885-023) supplemented with 10% (v/v) heat-inactivated fetal bovine serum (FBS; Gibco #10500064) and 1% (v/v) Penicillin-Streptomycin (Gibco #15140-122). At 80%–90% confluency, cells were washed with Dulbecco's Phosphate-Buffered Saline (DPBS; Gibco #14190250), dissociated using TrypLE Select (Gibco #A1217701), and reseeded in fresh complete growth medium in T75 flasks (Thermo Scientific #156499).

2.3. Cell seeding and compound treatment

All test compounds were provided as DMSO solutions and diluted to stock concentrations using a 10-fold serial dilution. Phenotypic reference compounds that are known to induce a distinctive phenotype in a variety of cells [35], including Etoposide (Sigma-Aldrich #E1383), Fenbendazole (Sigma-Aldrich #F5396), and Staurosporine (Sigma-Aldrich #19-123-M), were purchased from Sigma-Aldrich and prepared as 10 mM stocks in DMSO (Sigma-Aldrich #D8418). Assay-ready plates containing test compounds, phenotypic reference compounds, and DMSO (negative controls) were prepared using an I-DOT (immediate drop-on-demand technology) liquid handler. Compounds were dispensed in precise volumes to achieve final treatment concentrations of 0.1, 0.3, 1, 3, 5, and 10 μM for test compounds; 3 μM for the positive controls Etoposide and Fenbendazole; 0.3 μM for the Staurosporine control; and 0.1% for DMSO.

Each of the 8 compound plates included 20 or 21 test compounds at 6 concentrations and 2 technical replicates, 3 phenotypic reference compounds at 1 concentration and 2 technical replicates, and 14 wells of DMSO. The distribution of treatments and controls on the compound plates was optimized using PLAID (Plate Layouts using Artificial Intelligence Design) [36], a flexible constraint-programming model that ensures a robust microplate layout.

Prior to cell seeding on the assay ready plates, the compounds pre-spotted onto the assay plates were dissolved by adding 20 μL of complete growth medium per well using a Biotek Multiflo FX dispenser. The plates were then placed on a shaker at 150 rpm at room temperature for one hour to ensure complete dissolution of the compounds. Cells from passage 24 were counted and seeded on top of the plates with dissolved compounds at a density of 1100 cells per well by dispensing 20 μL/well of cell suspension with Biotek Multiflo FX dispenser, resulting in a total final volume of 40 μl/well. The seeded plates were left to stand at room temperature for 20 min before being incubated at 37 °C in a 5% CO_2 atmosphere for 48 h.

2.4. Cell painting assay

Cell painting experiments were conducted according to the protocol described [37] with minor modifications. The fluorescent labeling process was performed in an automated laboratory setup, which included a UR10 robotic arm for handling plates, a Biotek MultiFlo FX for dispensing solutions, and a BlueWasher (Blue Cat Bio) for washing steps. In brief, MCF7 cells were cultured in 384-well plates and treated with compounds at various doses for 48 h. Following treatment, cells were live-cell-labeled with MitoTracker Deep Red (Invitrogen) and subsequently fixed with 4% paraformaldehyde (Histolab) for 20 min. After fixation, cells were permeabilized with 0.1% Triton X-100 (Cytiva) for 20 min and stained with a solution containing Hoechst 33342, Wheat Germ Agglutinin, Phalloidin, SYTO 14, and Concanavalin A (all from Thermo Fisher Scientific) for 20 min. Finally, plates were washed, sealed, and stored at 4 °C, protected from light until image acquisition.

2.5. Imaging and image pre-processing

Fluorescence microscopy was performed using a high-throughput ImageXpress Micro XLS (Molecular Devices), capturing nine sites per well across five fluorescence channels. Excitation and emission wavelengths were optimized for each dye. Image processing and analysis were conducted with CellProfiler and CellPose 2.0 [38], including

quality control, illumination correction, segmentation, and feature extraction. Quality control was performed via filtering out images that surpassed intensity limits in selected features measuring key intensity levels of segmented objects. Illumination correction for CellProfiler and DINO images was calculated and applied via the CellProfiler *CorrectIlluminationApply* module using the *FitPolynomial* function. For DeepProfiler, the illumination correction function inside the DeepProfiler package was used according to [39]. A total of 107,405 images were processed. This number comprises of 8 384-well plates that were imaged at 9 field of views ("sites") per well in five-channels. Outer wells were left out and QC-flagged images removed, which results in the above number. Segmentation results were saved as binary masks and later used as input for feature extraction (CellProfiler), image cropping (DeepProfiler) and training (DINO).

2.6. CellProfiler

A feature extraction pipeline using Cell Profiler version 4.3.4 (available at https://doi.org/10.17044/scilifelab.21378906.v2 and described in [10]) was run and resulted in a 2160-dimensional feature vector for each cell. These features capture fluorescence intensity, location, granularity, and other measurements of the imaged cells, grouped into cell, nucleus, and cytoplasm features.

2.7. DeepProfiler

DeepProfiler is a convolutional neural network trained in a weakly supervised setting, as described by [15]. In this study, we utilized the pre-trained model CellPainting_CNN with weights from [15] for inference without additional training. Features were extracted as 672-dimensional vectors from the block_6a_activation layer of the EfficientNetB0 architecture, generating one feature vector per single-cell crop. To ensure accurate representation, we used nucleus-centered image crops with a resolution of 150 × 150 pixels. The cropping size was selected to fully encapsulate 95% of all cells. Illumination correction was applied during the DeepProfiler feature extraction process, using a down_scale_factor of 8 and a median_filter_size of 48. Importantly, segmentation masks were omitted during inference to retain contextual neighborhood information surrounding the cells.

2.8. DINO

Self-distillation with no labels (DINO) falls under the self-supervised learning paradigm to learn representations from images without the need for explicit label annotation [16]. The approach combines identical student and teacher networks to learn image features through self-distillation. We trained DINO with a ViT-small network, a patch size of 16 pixels, and five input channels. We adapted the architecture by [19], to allow for multi-channel inputs. The network was trained from scratch with a stratified and sub-sampled subset of our Cell Death dataset consisting of > 185,000 single-cell images, comprising all compounds at all concentrations (including DMSO). As network inputs, we used the segmented five-channel site-level images, center-cropped around the nuclei, and stacked the images into a (150 x 150 x 5) image for each cell. We excluded cells at the borders (with a 150 pixel radius) to avoid cells being cut off. For training purposes, pixel intensities were normalized to [0, 1] and images resized to 224 x 224 pixels. Using the same set of augmentations as [19] (random channel dropping, random brightness and contrast, artificial warping), we trained the DINO network for 200 epochs with a learning rate of 0.0005. Training was performed on a single NVIDIA GTX 3090 with a batch size of 20 for 3 days. Inference was performed on the teacher network, resulting in 384-dimensional single-cell feature vectors. To evaluate the model, we inferred 625,000 stratified, randomly sampled cells from our cell death dataset that were held out during training. A full set of hyperparameters can be found in Table S3. For a full description of the augmentations, we refer to the *config_finetune_unmasked.yaml* on our FigShare.

2.9. Data post-processing

To ensure comparability, we applied the same post-processing and analysis pipeline to all three feature sets. In the following, we will also refer to the extracted feature vectors as "profiles". Feature vectors were first standardized to the negative control (DMSO) by subtracting the mean and dividing by the standard deviation of the DMSO cells per plate. For our aggregated analysis, we calculated the median profile per field of view ("site"). Feature selection was performed by removing missing value features, highly correlated (> 99%), low variance (< 0.01%) and blacklisted features as defined in the pycytominer package [40]. Additionally, features with > 25% constant values across cells were removed. This resulted in just over 1400 features for CellProfiler, 384 features for DINO, and 672 for DeepProfiler, respectively. Additionally, we excluded cells bordering the images within a 150 pixel radius to avoid cells being cut off. (this was done post feature extraction to ensure correct performance of the neural network-based feature extractors). In short, we applied this filter: $150 \leq nuclei_{location_{x,y}} \leq 2850$. A summary of cell counts per compound concentration prior to filtering and sampling can be found in Supplemental Figure S1 as well as viability plots in Supplemental Figures S2-S4.

2.10. Sampling

In order to obtain cell sets more representative of the true population distribution and taking into account the over-abundance of DMSO wells in our plate layouts, we implemented a sampling strategy:

1. Perform a stratified sampling by plate, well, and site (the latter iff the resulting cell count after sampling is large enough).
2. To obtain the specific sampling per well/ site follow the strategy below, which we denote as well_ratio sampling:
   Let $s \in (0, 1]$ be a user-defined sampling rate and $n \in \mathbb{N}$ a predefined minimum number of cells per condition. Define W as the set of all wells belonging to DMSO and V_U as all wells belonging to Treatment U (T be the set of all treatments). Then, $D_w := \{d_i \mid d_i \in N \wedge d_i \in w, w \in W\}$ describes the set of negative control cells of a specific well W and $U_v := \{t_i \mid t_i \in C_U \wedge t_i \in v, v \in V_U\}\}$ the treated cells of a specific well V and compound U. Then, we define well-ratio sampling as:

$$w_{\text{sampling}} = \begin{cases} s, & \text{if } U \neq \text{DMSO} \\ r, & \text{otherwise} \end{cases}$$
$$r = \frac{\frac{1}{|W|}\sum_{w \in \mathcal{W}} D_w}{\max_{t \in T}\left(\frac{1}{|\mathcal{V}_t|}\sum_{v \in \mathcal{V}_U} U_v\right)} \cdot s \tag{1}$$

   Here $w_{\text{sampling}} \in (0, 1]$ is the per-well sampling *fraction*: treated wells are sampled at rate s, while DMSO wells are down-weighted by the factor r, i.e. the ratio of the average DMSO cell count per well to the maximum average cell count per well across all treatments (scaled by s). The number of cells drawn from a well is $\lceil w_{\text{sampling}} \cdot c \rceil$ for a well containing c cells, and a minimum of n cells per condition was retained where available. In case of $r \geq 1$, the sampling simplifies to:

$$w_{\text{sampling}} = \begin{cases} s, & \text{if } U == \text{DMSO} \\ 1, & \text{otherwise} \end{cases} \tag{2}$$

This sampling strategy was used in our unsupervised and exploratory analysis tasks as well as to define a training and inference set for our DINO training. In the unsupervised scenario, we matched our CellProfiler and DeepProfiler sets after etest filtering to match cell counts in the DINO set.

2.11. Energy distance

Energy distance (e-distance) [3] was originally implemented for single-cell sequencing data as a metric for perturbation strength. We calculated e-distances per compound concentration set using single-cell profiles and the first 50 principal components.

Let $x_1, \ldots, x_N \in C_1$ and $y_1, \ldots, y_M \in C_2$ be single-cell profiles, where N, M represent the respective sample size and C_T be our treatment spaces. We further define:

$$\delta_{XY} = \frac{1}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} \left\| x_i - y_j \right\|$$
$$\sigma_X = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \left\| x_i - x_j \right\| \tag{3}$$

The energy distance is then defined as:

$$E(X,Y) := 2\delta_{XY} - \sigma_X - \sigma_Y \tag{4}$$

Using the e-distance as a test statistic, we performed a permutation test with 50,000 permutations to calculate a p-value for each compound concentration in relation to the negative control. We removed all compound concentrations with an adjusted p-value > 0.05. For multiple testing correction, we used the *holm-sidak* approach.

2.12. Grit score

Grit score measures the perturbation strength and reproducibility of a given treatment group. We calculated grit scores on a compound-concentration level using aggregated profiles. We filtered compound concentrations with a grit score below 1.5.

Let $d_1, \ldots, d_N \in N$ be morphological profiles of a control group and $t \in C_T$ a target profile of given treatment T with M replicate profiles $r_1, \ldots, r_M$. Then define $D := \{d_1, \ldots, d_M\}$ where $d_i = corr(t, c_i)$ and $S := \{s_1, \ldots, s_M\}$ where $s_i = corr(t, r_i)$. Let further μ_D, σ_D be the respective statistics of D. Then the grit score G is defined as:

$$g_i = \frac{s_i - \mu_D}{\sigma_D}$$
$$G = \frac{1}{M} \sum_{i=1}^{M} g_i \tag{5}$$

where M represents the number of replicates.

Both e-distance and grit score were calculated with respect to the negative control, which was then omitted for further analysis. Reference compounds in Supplemental Table S2 that were not annotated with specific cell death subtypes were omitted.

2.13. Mean average precision

Mean average precision (mAP) is a performance metric used in cellular profiling to assess how well a ranking method brings perturbation groups to the top of a list. It combines precision and recall into a single scalar averaging the precision values at every rank where a relevant item appears, and then averages those values across all queries [41]. We computed mAP across perturbations by treating each unique compound concentration as a separate class, with profiles from sites treated with the same compound concentration. For each query, we calculated the cosine similarity between its profile and all others, ranked the profiles by descending similarity, recorded precision at each rank where a relevant profile occurred (the number of relevant profiles in the top k divided by k) and averaged those precision values to obtain the average precision (AP). We averaged AP over all queries to yield the mAP. For statistical evaluation, we selected a *null_size* of *n = 500,000*, referring to the Supplemental materials by Kalinin et al. [41]. We selected a p-value cut-off of 0.05 as well as *holm-sidak* for multiple testing correction (we selected these parameters to ensure comparability with our etest calculation).

2.14. Visualization and clustering

We used the scanpy [42] library to calculate PCA, neighborhood graphs, and UMAP embeddings. We trained our UMAP on single-cell profiles using PAGA for initialization and cosine similarity as the distance metric. For visualization purposes, cells extracted with CellProfiler and DeepProfiler were sampled to 10% in a stratified manner. For DINO results, we used the full data set post-filtering as it matched the sampled CellProfiler and DeepProfiler sets in size.

2.15. k^* distribution analysis

To compare our feature representations, we computed k^* distributions for each feature extractor (CellProfiler, DeepProfiler, and DINO) at both the aggregated and single-cell level. The k^* statistic captures local neighborhood consistency with respect to class membership and was proposed as a way to compare latent spaces [43]. Given a sample x_i with true label y_i, we compute pairwise cosine distances between all samples and rank them in ascending order. The k^* value is defined as the index of the first nearest neighbor whose class label differs from that of x_i, formally expressed as

$$k_i^* = \min \left\{ k \in \mathbb{N} : y_{\pi_k(i)} \neq y_i \right\},$$

where $\pi_k(i)$ denotes the index of the kth nearest neighbor of x_i, and $y_{\pi_k(i)}$ is its associated class label.

To account for class imbalance, k_i^* is normalized by the total number of samples n_{y_i} belonging to class y_i, yielding the normalized score:

$$\tilde{k}_i^* = \frac{k_i^*}{n_{y_i}}.$$

The resulting set of values $\tilde{k}_i^*$ for all samples $i \in S$ of a given class defines the k^* distribution for that class. We then summarize this distribution using three descriptive statistics. The mean of the distribution

$$\mu_{k^*} = \frac{1}{|S|} \sum_{i \in S} \tilde{k}_i^*,$$

the standard deviation

$$\sigma_{k^*} = \sqrt{\frac{1}{|S|} \sum_{i \in S} (\tilde{k}_i^* - \mu_{k^*})^2},$$

and the skewness defined as

$$\gamma_{k^*} = \frac{\frac{1}{|S|} \sum_{i \in S} (\tilde{k}_i^* - \mu_{k^*})^3}{\left(\frac{1}{|S|} \sum_{i \in S} (\tilde{k}_i^* - \mu_{k^*})^2 \right)^{3/2}}.$$

We report these statistics to characterize the local neighborhood structure in feature space. A lower μ_{k^*} indicates that class members tend to be surrounded by similar samples, while a higher γ_{k^*} reflects skewness toward tight clustering. Conversely, negative skewness suggests fragmented or overlapping class neighborhoods, where a γ_{k^*} <-0.5 indicates clustered distributions, a values of $-0.5 < \gamma_{k^*} < 0.5$ indicates overlapped and $\gamma_{k^*} > 0.5$ fractured distributions [43].

2.16. Subphenotype reproducibility and target separation

To test whether morphological subpopulations within a class are reproducible independently of the two-dimensional UMAP projection, we clustered cells directly in the original feature space using Leiden community detection on a cosine k-nearest-neighbor graph, and quantified reproducibility by re-clustering bootstrap resamples of the cells and computing the adjusted Rand index (ARI) between each resample and a reference clustering. A feature-shuffled null, in which each feature was permuted independently across cells to remove covariance structure, provided a no-structure baseline, and DMSO control cells served as a biological comparison. Separation of compound-target groups within

Table 1
Number of samples used for aggregated and Single-Cell classification models across Feature Extractors.

Cell death subtype (label)	CellProfiler		DeepProfiler		DINO	
	Aggregated	Single cells	Aggregated	Single cells	Aggregated	Single cells
Apoptosis (0)	2375	65,000	2371	65,000	1937	65,272
Autophagy (1)	598	30,000	860	30,000	666	32,167
Ferroptosis (2)	617	30,000	843	30,000	662	28,223
Immunogenic Cell Death (3)	1181	32,000	1078	32,000	1091	29,472
Necroptosis (4)	515	20,000	535	20,000	505	18,468
Pyroptosis (5)	853	33,000	848	33,000	819	32,992
Total	6139	210,000	6535	210,000	5680	206,594

the apoptosis class was assessed in the same feature space using three complementary statistics: the silhouette coefficient (cosine distance) computed on target labels; the k^* neighborhood-consistency statistic described above, with a target-label permutation test; and pairwise energy distance (Section 2.11) with a permutation test and Holm–Šidák correction across target-group pairs.

2.17. Supervised model training

The extracted profiles from CellProfiler, DeepProfiler, and DINO were used to train a classifier both on single-cell profiles and the aggregated profiles. First, profiles were split into a training and testing set (70/30), stratified over Cell Death subtypes. As our dataset contained a limited number of compounds per Cell Death subtype, we implemented a split with no overlap between seeding wells between the training and testing set. As this well-level split holds out whole wells but not whole compounds, we additionally evaluated generalization with a leave-compound-out split, in which all wells of a compound, across all concentrations, were assigned exclusively to either the training or the test set. A single fixed, class-balanced gradient-boosted classifier (LightGBM with extra-trees) was retrained under this split and, for comparison, under the original well-level split, with per-class scores reported alongside (Supplemental Table S4, Supplemental Figure S24). To account for the non-independence of single cells, we additionally report bootstrap 95% confidence intervals for the macro-F1, resampling whole wells or whole compounds — rather than individual cells — with replacement. Next, a set of models was trained using the AutoGluon Tabular framework [44]. The model selection included models from the XGBoost, CatBoost, Extra Trees, Random Forest, and FastAI NN families [45–48]. AutoGluon Tabular runs an automated stacking, combination, cross-validation and hyper-parameter tuning of models and reports a ranked list of trained models. We ran AutoGluon Tabular on the "high_quality" setting without custom models and removed the KNN model family. The training time was set to 24 h. The two top-performing models were selected for reporting, ensuring overlap between the training scenarios. In addition to single-cell and aggregate classifications, we implemented a third strategy ('Majority vote') where we first predicted cell death subtypes on single-cell level followed by a majority vote over predicted labels per site. A prediction was assigned as correct if and only if the majority-voted label matched the true label. To evaluate our classification model, we reported F1, Recall, Precision as well as Accuracy. All metrics were reported on macro and micro (weighted) level. Additionally, we generated confusion matrices for all models.

2.18. Data and code availability

The code to run the analysis and generate the figures can be found on Github (https://github.com/pharmbio/sc_celldeath). The feature profiles, metrics, and splits to run the analysis and supervised training can be found on FigShare: https://doi.org/10.17044/scilifelab.28202864.

3. Results

3.1. Measuring perturbation strength using e-distance enables robust cell filtering

Before quantifying perturbation effects, we first evaluated cell viability and cell counts to determine whether dose-dependent responses were detectable and to exclude compounds or conditions that showed no activity or induced widespread cell death (Fig. 1D, S1, S2, S3, S4). We then calculated grit scores and e-distances for groups of single cells treated with the same compound concentration, including only those groups with at least 200 cells. Dimethyl Sulfoxide (DMSO), used as a solvent control, served as the negative control in these assessments. For all downstream analyses, DMSO-treated cells were excluded, as they did not exhibit phenotypes consistent with any of the programmed cell death subtypes under investigation.

Grit calculations resulted in 54.04% of sets kept for CellProfiler features, 53.43% for DeepProfiler features, and 51.56% for DINO features. Notably, four compound concentration sets extracted with CellProfiler showed significantly higher grit scores than in DeepProfiler. Further inspection revealed that these sets did not appear in DeepProfiler features as post-processing of the cells resulted in lower cell counts and subsequent removal (Table 1). The Venn diagram (Fig. 1C) shows that 134 sets resulted in a grit > 1.5 for all feature extractors, with 20 sets only in DINO and CellProfiler features. Inspection of the distributions of grit scores further reveals that DINO features result in a significant proportion of sets with grit scores > 4. This indicates strong phenotypical changes when compared to DMSO (Supplemental Figure S5).

Etest and e-distance results revealed similar distributions in distance and significance results between CellProfiler, DeepProfiler and DINO features (Supplemental Figure S6). For CellProfiler, higher absolute distances were observed with a maximum e-distance of 11,590 as compared to 3090 in DeepProfiler and 420.93 for DINO (Supplemental Figure S6). Due to the non-interpretability of the distances, filtering was performed using results from the corresponding etest. Results show 74.90% of sets as significant for DeepProfiler features, 66.54% significant for CellProfiler features and 71.00% for DINO. Here, 155 common sets were observed between the three feature sets (Fig. 1C). It is noteworthy that the number of sets differs between grit and e-distance calculations, as the metrics have different underlying assumptions. Therefore, the intersection of significant etest and grit-scores > 1.5 was used as a cut-off for further analysis.

We further calculated mean average precisions for all compound concentrations, observing similar performance to the etest calculations with an overlap of 162 groups with etest in CellProfiler (Supplemental Figure S7A), 169 groups in DeepProfiler (Supplemental Figure S7B) and 187 groups with DINO (Supplemental Figure S7C). Additionally, we found similar levels of phenotypic activity between DeepProfiler (70% mAP, 74% etest), Supplemental Figure S7B) and DINO (79% mAP, 85% etest, Supplemental Figure S7C), with a larger difference in CellProfiler (81% mAP, 66% etest, Supplemental Figure S7A).

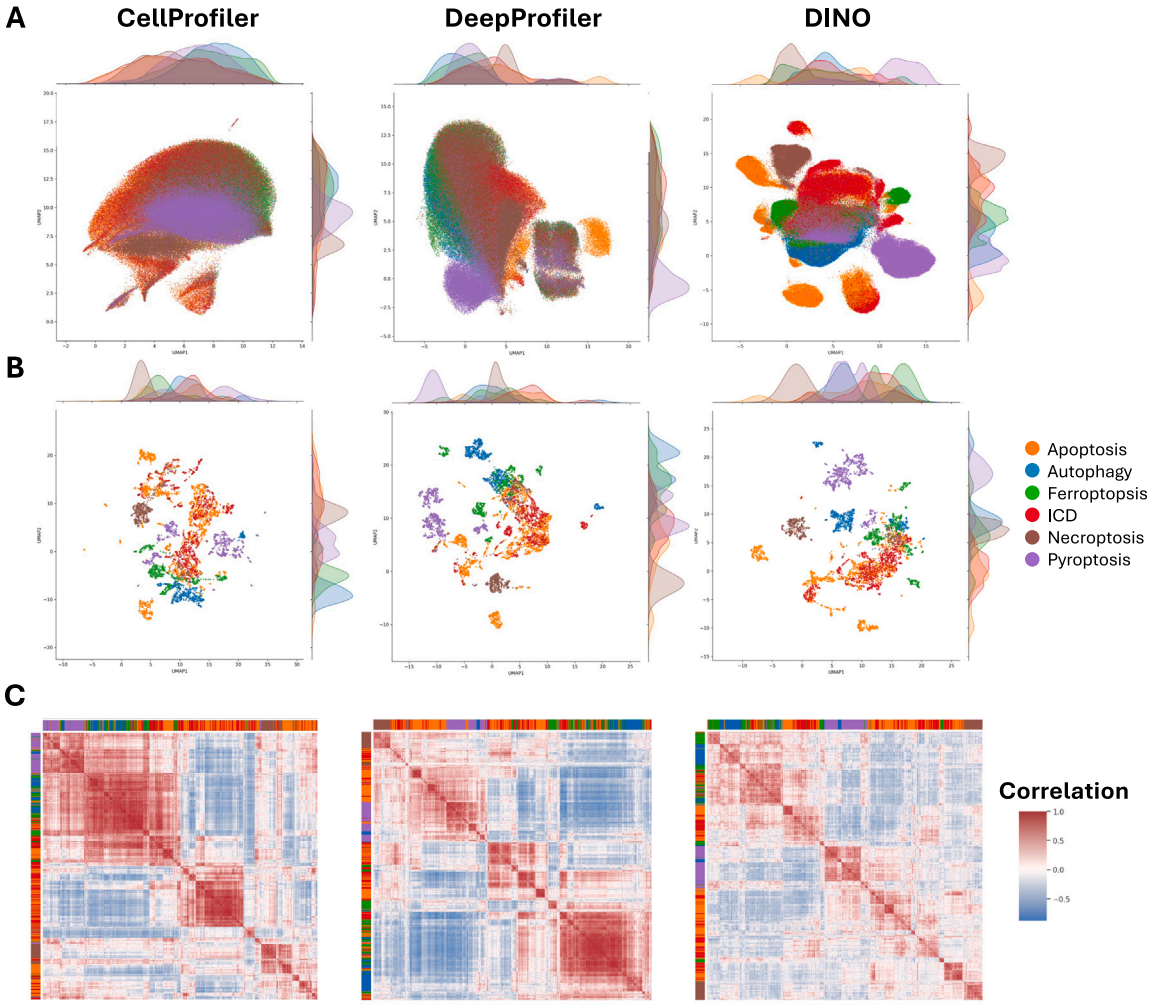

Fig. 2. Comparison of morphological profiling across feature extraction methods. (A) UMAP embeddings of cell morphologies extracted using CellProfiler, DeepProfiler, and DINO. Each point represents a single-cell, colored by its cell death subtype. Marginal density plots along the axes display the distribution of cells for each subtype, highlighting clustering patterns unique to each method. (B) UMAP embeddings of aggregated cell morphologies extracted using CellProfiler, DeepProfiler, and DINO. Each point represents a aggregated profile, colored by its cell death subtype. Marginal density plots along the axes display the distribution of cells for each subtype, highlighting clustering patterns unique to each method. (C) Correlation heatmaps comparing pairwise feature similarities among cell treatments. Rows and columns correspond to site-level aggregated profiles, with colors representing correlation values (red for positive and blue for negative correlations). The distribution of cell death subtypes is indicated by colored bars along the heatmap axes.

3.2. DINO features learn higher resolution morphological patterns on single-cell level

In order to characterize the selected PCD subtypes from a single-cell perspective, we used different feature extraction methods applied to single-cell data as well as aggregated (median cell) data.

We first compared the feature sets based on lower-dimensional UMAP embeddings. Our results show stronger separation of PCD subtypes at the single-cell level using DeepProfiler features compared to CellProfiler (Fig. 2A). A clear separation of the Pyroptosis class is observable in both the single-cell UMAP along with the marginal kernel density estimations, as well as in the corresponding correlation matrix (Fig. 2C). For CellProfiler features, we observe a positive correlation of profiles from the Pyroptosis, Ferroptosis, and Autophagy classes. For DeepProfiler features, we notice positive correlations between profiles from the Ferroptosis and Autophagy classes, but a clear separation from the Pyroptosis class. Further, a clear cluster formation of the Apoptosis class can be observed in the DeepProfiler embedding. We identified the compound FK866 to be responsible for the formation of this cluster as indicated in Supplemental Figure S8B middle. When comparing the results with the DINO embedding, more noticeable cluster formation on a single-cell level as well as in the correlation maps can be found (Fig. 2A and C DINO). Clusters appear to be more dense as indicated by the marginal KDEs in the single-cell UMAP as well as by highly correlated profiles in the corresponding heatmaps (Fig. 2). When taking into account the compounds, we observed a clear pattern; individual compounds corresponding to the observed clusters in morphological space (Supplemental Figure S8B). Density estimations of the PCD subtypes in our UMAP spaces revealed an overall more uniform distribution of cells in both CellProfiler and DeepProfiler features (Supplemental Figure S9A and B), while the subtypes in our DINO embedding space form multiple sub-clusters (Supplemental Figure S9C).

Subsequent analysis of median profiles revealed a clearer separation of PCD subtypes and specific compounds in all three feature sets (Fig. 1B and Supplemental Figure S8A). In all feature sets, we observed a region of strong overlap between the cell death subtypes and identified these regions as groups with lower grit and e-distance scores and less separable (Fig. 2B). Overall, the differences between feature sets on an aggregated level are less apparent when compared to single-cell level.

One noticeable observation in the DeepProfiler features is the clearly separated cluster in the UMAP embedding (Fig. 2A.) The cluster

Table 2
Summary of k^* distribution statistics for each cell death subtype, feature extractor, and level of representation. Lower μ_{k^*} and higher γ_{k^*} indicate tighter and more homogeneous class-specific neighborhoods.

Class	Extractor	Aggregated			Single-cell		
		μ_{k^*}	σ_{k^*}	γ_{k^*}	μ_{k^*}	σ_{k^*}	γ_{k^*}
Apoptosis	CP	0.017	0.022	1.47	0.001	0.001	4.90
	DP	0.023	0.031	1.43	0.002	0.007	4.91
	DINO	0.027	0.035	1.19	0.009	0.019	2.40
Autophagy	CP	0.124	0.104	0.68	0.001	0.001	3.29
	DP	0.145	0.145	0.63	0.001	0.002	5.50
	DINO	0.116	0.110	0.97	0.005	0.009	3.40
Ferroptosis	CP	0.063	0.055	1.10	0.001	0.001	6.26
	DP	0.061	0.064	1.16	0.001	0.001	5.41
	DINO	0.048	0.041	1.04	0.004	0.008	3.72
ICD	CP	0.011	0.013	1.80	0.001	0.000	4.21
	DP	0.017	0.023	1.80	0.001	0.001	6.38
	DINO	0.011	0.013	1.69	0.006	0.015	3.51
Necroptosis	CP	0.154	0.107	0.06	0.002	0.002	3.88
	DP	0.253	0.198	0.05	0.002	0.003	3.58
	DINO	0.329	0.243	−0.14	0.032	0.046	2.05
Pyroptosis	CP	0.237	0.177	0.45	0.002	0.003	4.26
	DP	0.446	0.248	−0.54	0.013	0.020	2.96
	DINO	0.286	0.183	0.16	0.059	0.071	1.53

was identified to consist of compounds from all PCD subtypes without any distinct pattern indicative of its formation (Fig. 2 DeepProfiler and Supplemental Figure S10). Analysis of the corresponding microscopy images using our *CellViewer* tool did not reveal any patterns in the images that would explain this cluster (Supplemental Figure S10C). Further analysis of the original feature intensities revealed a distinct pattern in all cells originating from that cluster (Supplemental Figure S10B), but due to the non-interpretability of the features, we were limited in our understanding of this anomaly. As a removal of the blob did not result in a change of the remaining morphological space, we kept the points in for all further analysis.

To further quantify the differences in the embeddings, we calculated the corresponding k^* distributions S11. The summary statistics reveal differences in the distributions between single-cell and aggregated profiles across all feature extractors, where overall skewness is higher in single-cell profiles (Table 2). While all single-cell distributions are positively skewed, the degree of skewness varies between cell death subtypes and feature extractors. Overall, we observe lower skewness values for DINO features (1.53–3.73) compared to CellProfiler (3.29–6.21) and DeepProfiler (2.96–6.38). Similarly, the highest degree of skewness is found in the apoptosis, ferroptosis and ICD class, while the remaining classes show overall lower skewness. For aggregated profiles, we observed a shift in distributions, specifically in the necroptosis, pyropotosis and autophagy class. For nectroptosis, we see the emergence of overlapped skewness patterns for all feature extractors, while for pyroptosis our DeepProfiler features indicate a clustered distribution as indicated by a skweness below −0.5.

3.3. Single-cell density analysis reveals biological heterogeneity within apoptotic cells

Building on the previous results, we sought to explore the heterogeneity within our PCD subtypes. For this, we focused on the DINO-derived features. We selected apoptosis for further analysis due to its high variability across compounds, providing a case for studying the cellular heterogeneity.

To investigate, a new UMAP embedding was trained on the single-cell DINO features of the apoptotic cells. The embedding revealed distinct cluster formations within the morphological space (Fig. 3 and Supplemental Figure S12). To verify that these subpopulations are not an artifact of the two-dimensional embedding, we clustered the apoptotic cells directly in the original 384-dimensional DINO feature space and assessed their reproducibility by bootstrap resampling: the clustering was highly stable (adjusted Rand index 0.80 ± 0.04 across resamples), far above a feature-shuffled null (0.01) and higher than DMSO control cells (0.54), confirming that the subpopulations reflect reproducible morphological structure rather than projection distortion (Supplemental Figure S25). To further understand the emergence of these clusters, we utilized our *CellViewer* tool to examine different regions within the UMAP. Representative cells from each region were identified using a k-medoid algorithm, and their corresponding cell images were annotated directly onto the UMAP (Fig. 3A). The analysis revealed that specific regions in the morphological space were associated with distinct cellular morphologies. Notably, cells in the region highlighted in the middle exhibited morphological traits typically associated with apoptosis, including shrinkage, isolation, nuclear fragmentation, and vesicle formation. Interestingly, this region contained cells originating from a diverse array of compounds (Supplemental Figure S12). Outward from this central apoptotic region, more distinct morphological patterns emerged, highlighting morphological heterogeneity within the apoptosis subtype.

Given this observed heterogeneity, we next investigated its potential origins. As outlined, cells in this experiment were treated with a wide range of compounds known to induce apoptosis. To elucidate the relationship between morphology and molecular mechanisms, compounds were next annotated with their respective targets and which were mapped onto the UMAP. This revealed a connection between distinct regions of the UMAP and specific molecular targets (Fig. 3B and Supplemental Figure S13C). To confirm that these morphological subphenotypes—and their correspondence to molecular targets—are not artifacts of the two-dimensional UMAP projection, we tested the target association directly in the original feature space. Within the apoptosis class, the four molecular-target groups with at least two compounds (Topoisomerase I, Topoisomerase II, DNA Polymerase, Bcl-2) were all significantly separated by energy distance (permutation test, Holm–Šidák $p < 0.05$; Supplemental Figure S22), and target labels showed neighborhood consistency (k^*) far exceeding a permuted-label null ($p < 0.001$), indicating that the subphenotypes carry genuine target-associated signal rather than projection artifacts. This organization was subtle but most pronounced for DINO features, which best resolved the subphenotypes and were the only representation with a positive target silhouette (Supplemental Figure S21). For instance, cells in the bottom-right region (Fig. 3A) were predominantly treated with compounds targeting Topoisomerases I/II, DNA polymerases, or intercalating DNA agents (Fig. 3B). The morphological characteristics in this

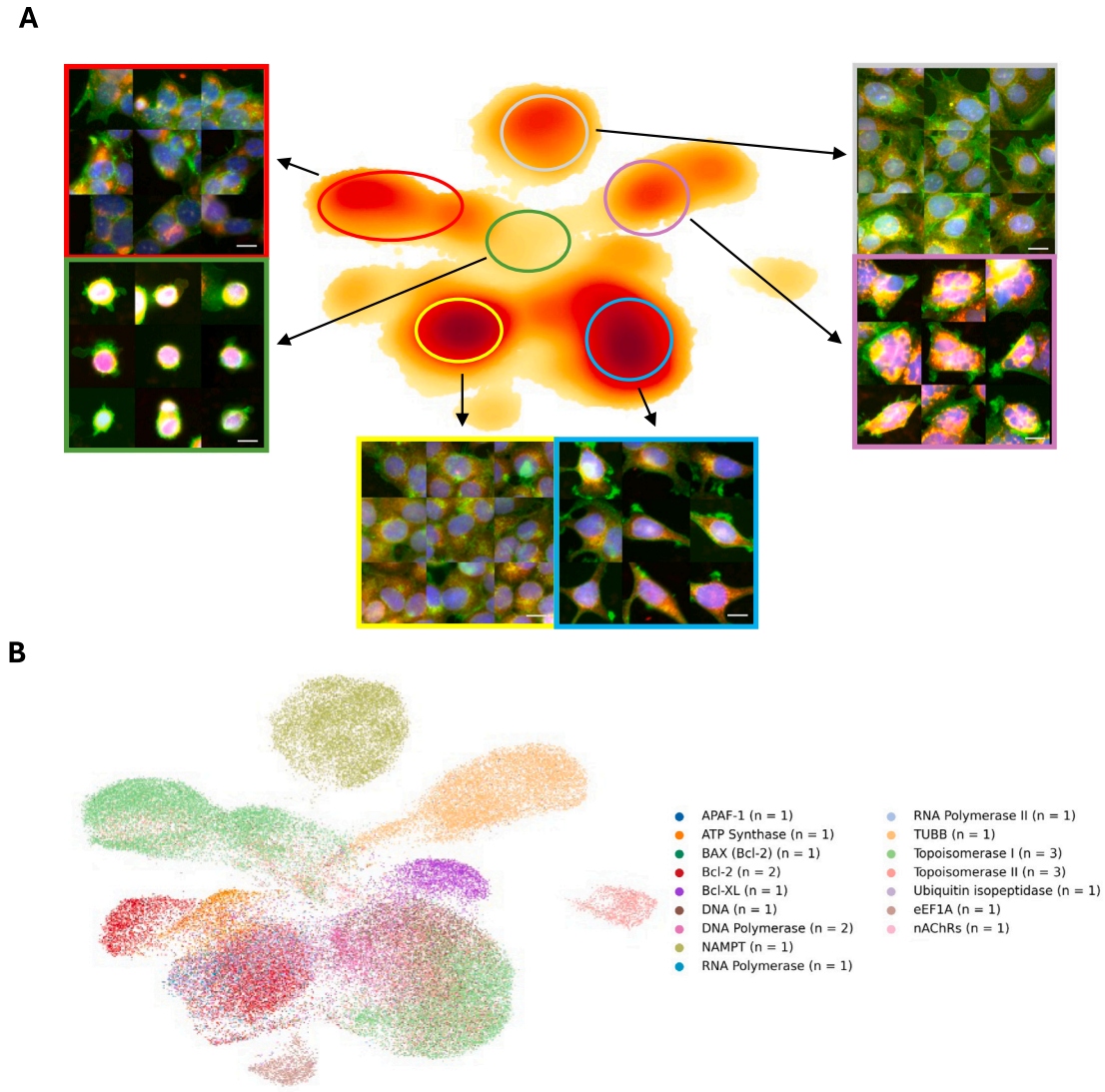

Fig. 3. Density mapping of cells from the apoptosis subtype on DINO embedding. Scale bar corresponds to 20 μm. (A) UMAP showing the distribution of cell morphologies across treatments in the apoptosis class, with warmer colors (yellow to red) indicating higher densities. Colored circles highlight specific clusters linked to distinct phenotypes. Representative cell images adjacent to the UMAP are overlays of the Hoechst (nucleus, blue), Mitotracker (red/yellow) and Phalloidin (Cytoskeleton/ actin, green) channels. (B) Single-cell UMAP embedding of the apoptosis class colored by their annotated targets and number of unique compounds per target. Specific target, compound matching can be found in Supplemental Table S1.

region reflect these molecular mechanisms specifically by changes in the nucleus channel (indicated by blue color).

Analogous analyses using DeepProfiler and CellProfiler features demonstrated similar trends in morphological subspaces, albeit with less pronounced cluster formation (Supplemental Figure S13A and B, Supplemental Table S1). Overall, DINO-derived features provided the clearest differentiation of sets of cells while maintaining the most homogeneity among distinct regions in the morphological space. These findings underscore the ability of single-cell morphological features to reveal fine-grained heterogeneity within PCD subtypes and highlight the superior performance of DINO features in distinguishing regions connected to a molecular target.

3.4. Single-cell analysis reveal dose-dependent morphological effects

We next investigated how compound concentration influences morphological changes on single-cell level. We projected chemical perturbations at varying concentrations onto the morphological space of the corresponding cell death subtypes. More precisely, we first filtered our datasets to the respective subtypes, learned a new UMAP embedding, and subsequently projected the compound concentrations into that space and calculated their density distributions. As a showcase, we focused on the apoptosis class, which contained the highest abundance of individual compounds. We randomly selected three compounds (Cladribine, SN 38, and Topotecan) to examine their dose-dependent morphologies. We were able to detect shifts in morphology within the morphological space, driven by compound dosage (Figs. 4, S14, S15). Notably, SN 38 exhibited the most prominent morphological changes, with cells shifting from a cluster in the bottom right to a region in the top left of the apoptosis-specific morphological space, based on DINO features (Fig. 4B). We used the *CellViewer* tool to identify key representative cells across concentrations. Representative cells highlighted this transition, showing a progression from triangular-shaped cells with large nuclei at the lowest concentration (0.1 μM) to elongated cells at the highest concentration (5.0 μM) as well as cytoskeletal re-organization. A similar dose-dependent effect was observed for Cladribine (Fig. 4C), where cells at the lowest concentration resembled untreated MCF-7 cells. With increasing concentrations, cells exhibited larger nuclei, stronger mitochondrial signals, and a lack

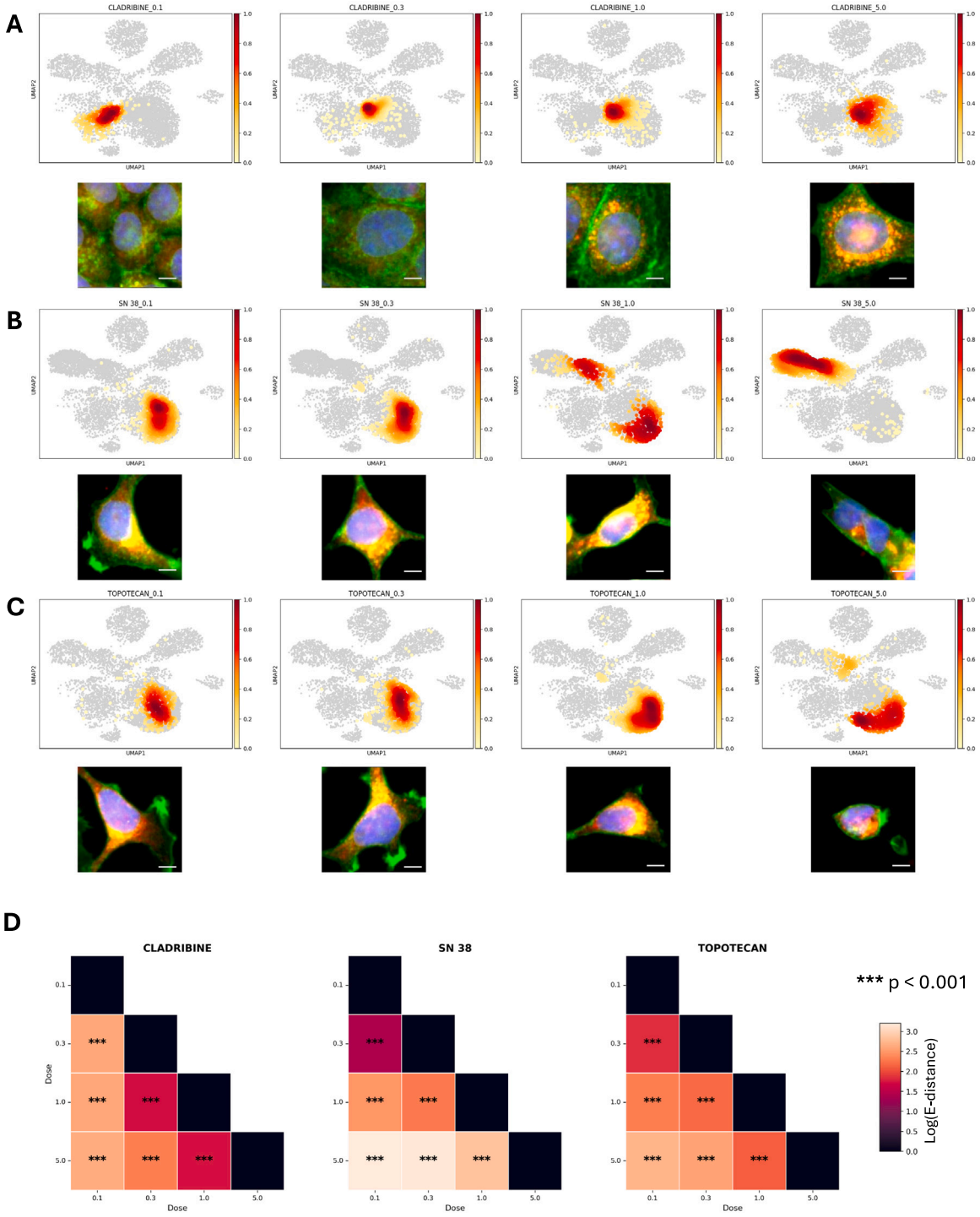

Fig. 4. UMAP representation of single-cell morphological profiles treated with varying concentrations of compounds, accompanied by representative cell images for the apoptosis class in DINO features. Scale bar corresponds to 10 μm. Each row shows UMAP projections of cells treated with increasing concentrations (0.1, 0.3, 1.0, 3.0, and 5.0 μ M). UMAP plots highlight the distribution of treated cells (colored) relative to the remaining cells in the apoptosis. Representative cell images below each UMAP depict an overlay of nuclei (blue), cytoskeleton (green), and mitochondria (red/yellow). Cells were identified with *CellViewer*. (A) UMAP projections and representative images of cells treated with Cladribine. (B) UMAP projections and representative images of cells treated with SN 38. (C) UMAP projections and representative images of cells treated with Topotecan. (D) Pair-wise energy distances (log) between concentrations (μ M) for Cladribine, SN 38 and Topotecan annotated with significance level calculated via etest.

of neighboring cells. However, not all compounds demonstrated pronounced changes. For instance, Topotecan showed lower morphological alterations at lower concentrations, with notable shifts only emerging at the highest tested concentration. We corroborated these observations by pair-wise e-testing between the respective compound concentrations (Figs. 4D, S14D, and S15D). Our results showed significant changes in morphology between all compound concentrations. Furthermore, energy distances confirmed previous observations where SN 38 presented

Table 3
Classification performance for the aggregated and site-level majority-vote (multiple-instance) strategies across feature sets. Columns give Accuracy, Precision, Recall and F1 (all macro); bold marks the best value per metric. Raw single-cell classification metrics are reported in Supplemental Table S5.

Model	Strategy	Accuracy	Precision	Recall	F1
CellProfiler (CP)					
LightGBMXT_{L2}	Majority vote	0.8075	0.8499	0.8407	0.8516
	Aggregated	0.8954	**0.9335**	**0.9001**	**0.9149**
$\text{NeuralNetFastAI}_{L1}$	Majority vote	0.8193	0.8645	0.8343	0.8441
	Aggregated	0.7865	0.8265	0.8028	0.8019
DeepProfiler (DP)					
LightGBMXT_{L2}	Majority vote	0.8186	0.8685	0.8000	0.8212
	Aggregated	**0.8997**	0.9260	0.8966	0.9000
$\text{NeuralNetFastAI}_{L1}$	Majority vote	0.8038	0.8487	0.7875	0.8052
	Aggregated	0.7986	0.8452	0.8017	0.8091
DINO					
LightGBMXT_{L2}	Majority vote	0.8556	0.8800	0.8723	0.8712
	Aggregated	0.8691	0.8870	0.8854	0.8851
$\text{NeuralNetFastAI}_{L1}$	Majority vote	0.8381	0.8701	0.8610	0.8572
	Aggregated	0.7050	0.7500	0.7456	0.7345

the most noticeable changes in morphology at higher concentrations as well as noticeable shifts between the concentrations of 0.1 and 0.3 μM in Cladribine.

To further explore how we could use dose–response effects in our understanding of PCD, we examined features derived from the CellProfiler suite. Specific cellular features have previously been used to investigate effects relevant to cell death and the cell cycle, such as the integrated intensity of the nucleus [6,49–51]. To demonstrate its utility, we analyzed a range of compounds and their corresponding feature distributions in the normalized integrated intensity across concentrations (Supplemental Figure S16). Our findings revealed clear dose-dependent integrated DNA intensity shift, indicating a change in cell cycle such as G1 or G2 arrests, a characteristic effect known to be induced by these families of compounds. These results underscore how morphological features, such as those derived from CellProfiler, can provide quantitative insights into dose-dependent cellular responses, leveraging information provided by single cells.

3.5. Aggregate profiles outperform single-cells in cell death classification

After identifying distinct cell death subtypes through unsupervised analysis, we proceeded to evaluate whether single-cell morphological profiles could be utilized to accurately classify cell death subtypes. Because single cells are noisy in isolation, we focused on two population-level strategies: classification of aggregated (median) profiles, and a site-level *majority-vote* strategy that treats each imaging site as a bag of single cells and assigns it the majority label of its cells—a simple multiple-instance learning (MIL) formulation. Raw single-cell classification is reported for completeness in Supplemental Table S5.

Classification of aggregated profiles performed strongly across all feature extractors. The LightGBMXT model trained on DINO features attained a macro F1 score of 88.51%, with a comparable performance to the model trained on CellProfiler features, which achieved a macro F1 score of 91.49% and the DeepProfiler model with an macro F1 score of 90.00% (Table 3).

As a second population-level strategy we applied the site-level majority-vote (MIL) approach, in which per-cell predictions from the single-cell classifier are pooled by majority vote within each site. This substantially improved on the raw single-cell classifiers (Supplemental Table S5). The LightGBMXT model trained on DINO features saw a performance boost of 7.26%, achieving an accuracy of 85.56% and a macro F1 score of 87.12% (Table 3). Similar performance deltas were observed across all remaining models, demonstrating the effectiveness of this aggregation strategy. However, despite the observed gains, none of the majority-voted models surpassed the top-performing models trained on aggregated profiles.

To further investigate model performance, we evaluated the confusion matrices for our classification models, showing the percentage of correctly predicted labels. Across all models, we observed a noticeable fraction of misclassifications between the ICD class (label 3) and apoptosis class (label 0) with up to 47.47% misclassification (Supplemental Figures S17-S19). The rate of misclassification was least pronounced in our DINO models, specifically in the single-cell models, where we observed a maximum of 21.09% misclassification between classes. As our feature extractions resulted in minimal differences in compound concentration composition and thereby train/test splitting, we further evaluated our models on an intersection of the test sets. We observed a slight increase in performance without change in the performance difference between the models as indicated by misclassification rates (Supplemental Figure S9).

Analysis of our compounds revealed structural homology between a few compounds in these classes; Etoposide (apoptosis) and Teniposide (ICD), as well as Doxorubicin/ Epirubicin (apoptosis) and Daunorubicin (ICD). We noticed similarities among their induced morphological effects, which could explain the misclassification of these specific classes (Supplemental Figure S9).

In an attempt to validate these results, we visualized the distribution of cell death classes in the UMAP spaces through density estimation (Supplemental Figure S9). We observed homogeneity among classes for both CellProfiler and DeepProfiler features (Supplemental Figure S9A and B), while DINO features exhibited significantly stronger spread, an observation we captured earlier through the k* distribution skewness (Table 2, Supplemental Figure S9C). We point out the overlap of the apoptosis and ICD classes in the UMAP spaces, further hinting at our observation of structurally similar compounds between the two classes (Supplemental Figure S9A-C).

4. Discussion

Effective handling of perturbations that produce little to no morphological change can be an essential tool in profiling studies. At the single-cell level, filtering out "unaffected" cells reduces noise and greatly enhances downstream detection of genuine morphological effects. In our implementation of the grit score, we noticed limitations in computing scores at the single-cell level for larger datasets, restricting its applicability in high-throughput experiments. For instance, on our cluster with 256 GB of RAM, grit scores could not be reliably computed for over 15,000 cells in one go, greatly curtailing its applicability in such experiments. To complement grit scores, we evaluated e-distance, a computationally more efficient metric suitable for large-scale single-cell datasets [3]. Our results showed that e-distance and grit scores produced comparable filtering outcomes, suggesting that either metric, or a combination, could be effectively used to identify treatments inducing significant morphological changes (Fig. 1C). The choice of filtering metric ultimately depends on the experimental context and

availability of computational resources, with e-distance offering scalability for single-cell data while grit scores provide robustness at the aggregated level. Additionally, we compared our results to the recently proposed mean average precision [41] and showed comparable results for DeepProfiler and DINO derived features (Supplemental Figure S7). For CellProfiler, we noticed a more pronounced difference in significant groups compared to etest. We leave a more thorough investigation of the effect of feature extractors on perturbation metrics open to future works.

Our unsupervised analysis illustrated that single-cell profiles offer a powerful means to study how different types of PCD emerge, revealing both subpopulation dynamics and distinct morphological characteristics (Figs. 2 and 3). A notable observation across all feature extractors was the non-homogeneity of PCD subtypes. Using the apoptosis class as an example, we were able to identify distinct cellular populations within the morphological space, each associated with specific morphological traits (Fig. 3). Indeed, apoptosis encompasses a wide range of intrinsic mechanisms [28,52]), leading to diverse morphological effects, such as the formation of apoptotic bodies, decrease in cellular volume, enlargement of the ER, and membrane blebbing or blistering, among others [28,53], which may explain the diversity of phenotypes that we identify.

To validate the identified apoptosis subpopulations, we mapped compound targets to these regions in the morphological space. Strikingly, this revealed a clear correlation between the subpopulations and specific target annotations, suggesting that distinct molecular mechanisms drive the observed morphological diversity (Fig. 3B). Compounds targeting similar pathways or molecular components clustered together within the morphological embedding, providing an additional layer of biological interpretation to the observed heterogeneity. Moreover, we demonstrated that dose-dependent changes in morphology further contribute to heterogeneity, as compounds induce distinct morphological effects only at specific concentration levels (Fig. 4 and Supplemental Figures S14 and S15) and certain cell cycle stages (Supplemental Figure S16).

The choice of feature extractor plays a crucial role in shaping the observed morphological patterns. Deep learning-based feature extraction methods, such as DeepProfiler and DINO, leverage fundamentally different architectures to learn features from imaging data. DeepProfiler employs a weakly supervised approach explicitly trained to differentiate treatments [15], while DINO uses a self-supervised paradigm to learn representations by minimizing a loss between teacher and student networks, focusing on local and global features [16]. Our results indicate that DINO tends to capture more intricate features, as evidenced by dense clustering of similar treatments in both UMAP embeddings and correlation maps (Figs. 2, 3 and Supplemental Figure S8). We note, however, that this is not a fully controlled architectural comparison: DINO was trained from scratch on our Cell Death dataset whereas DeepProfiler was used for inference with publicly released weights and no dataset-specific fine-tuning, so DINO's advantage may partly reflect dataset-specific adaptation rather than an intrinsic superiority of self-supervised over weakly-supervised representations.

As deep learning based features are inherently difficult to interpret biologically, it becomes necessary to use additional approaches like our mapping of underlying imaging data to validate their relevance. To further extend on this, explainability remains a critical consideration for deep learning approaches, especially in biological and medical applications [54]. While tools like CellProfiler provide interpretable features, neural network-based methods often require techniques like GradCAM [55] or visualization of attention heads in vision transformers [16] to understand what the model is focusing on during training and inference.

However, we point to anomalies, such as the unexplained cluster in the DeepProfiler features (Supplemental Figure S10), which suggest limitations in deep-learning approaches. We hypothesized that this cluster could originate from differences in the training set (using U2OS and A549 cell lines) compared to the MCF-7 cells we used for inference here in our study [15]. This underscores the importance of domain-specific fine-tuning when using larger pre-trained models and careful evaluation of obtained results.

Despite the strengths of single-cell profiles in uncovering subpopulation dynamics and morphological heterogeneity, several limitations must be acknowledged. While our supervised classification tasks achieved macro F1 scores of up to 79%, our single-cell classification models consistently underperformed compared to majority-voting and aggregated profile approaches. This performance gap may be attributed to an oversimplification of underlying population heterogeneity when relying solely on single-cell data and treating them as individual data points. It is further compounded by the pseudo-replicated nature of single-cell measurements: because cells from the same well and site are correlated, single-cell classification metrics are not drawn from statistically independent samples and should be interpreted as an upper bound on single-cell discriminative signal (Supplemental Figure S23). In an attempt to consider treatments more as populations, we implemented a simple majority voting multiple-instance learning strategy, reaching similar performance levels compared to the aggregated (Table 3). Recently, more advanced modeling approaches such as optimal transport, archetype analysis, and deep sets [2,56–58] have emerged, offering valuable insights into perturbation effects and cellular responses at the same time. These methods hold promise for future addressing the challenges associated with single-cell analysis by taking underlying distributions and their changes into consideration.

In addition to the aggregation levels, smaller performance differences were observed between CellProfiler, DeepProfiler, and DINO models, with disparities of up to 13% in accuracy (Table 3). In our pipeline, we matched dataset sizes as closely as possible to minimize scaling effects, as larger datasets have been shown to improve performance by better approximating the underlying data distributions [59,60]. Additionally, we assessed results on an intersection of our test sets ensuring equality between predicted sites. Here, we observed similar results, ensuring that model performance was not a consequence of differences in our data splitting approach (Supplemental Figure S9). Since aggregated results across feature extractors were consistent, we attribute the differences to the underlying architectures and specifically to DeepProfiler's training paradigm and the detected anomaly in its feature space (Supplemental Figure S10). Overall however, our findings are consistent with earlier studies showing that deep-learning-derived features (specifically from DINO networks) perform on par with or exceed CellProfiler features in classification tasks [13,15,20]. Our results underscore the importance of considering both training data size and feature extraction methodology when designing models for morphological profiling. They further highlight the potential trade-offs between capturing detailed biological variability and achieving state-of-the-art performance in classification tasks. Future works could explore strategies to enhance DINO's generalization capabilities, such as incorporating additional pre-training data, fine-tuning on class-specific labels, or using hybrid approaches that combine supervised and self-supervised learning paradigms or contrastive learning frameworks with other data modalities as seen in [61,62].

From a computational perspective, the scalability of single-cell morphological profiling remains a challenge. High-throughput analyses require efficient pre-processing and filtering strategies to ensure reliable downstream analyses. In this study, deep learning-based feature extraction methods not only matched or outperformed traditional tools like CellProfiler but also required fewer features to achieve these results. CellProfiler with extensive feature selection resulted in 1,400 hand-crafted features, while DeepProfiler and DINO had 672 and 384 features, respectively, to deliver comparable or better outcomes. Additionally, inference of features was obtainable in substantially less time with 2 h per plate in DeepProfiler and 1 h for DINO, both running on a single NVIDIA GTX 3090. As CellProfiler is, as of now, not capable of running feature extraction on GPUs, the inference times approached

> 4 h per plate. Overall, computational cost reductions significantly alleviate computational bottlenecks, making deep learning frameworks suitable for large-scale, high-throughput applications.

Finally, all analyses here use a single, well-characterized cell line (MCF7) as a model system to demonstrate the profiling workflow. MCF7-specific biology—for example its documented deficiency in caspase-3—will shape the particular cell-death morphologies observed. Accordingly, the biological observations we make are deliberately weak and exploratory, resting on literature-based compound annotations and observed treatment responses; the primary contribution of this work is the single-cell morphological-profiling methodology and the comparison of feature representations, and cell-line-general biological conclusions would require validation in additional models.

Single-cell morphological data can offer invaluable insights into cellular heterogeneity and subpopulation dynamics, complementing aggregated profiles for exploratory and hypothesis-driven analyses. However, in this work, we showed that for effective use, rigorous methodological choices are required, including feature extraction strategies and filtering metrics, tailored to the specific biological context. By integrating computationally efficient and scalable analysis methods, single-cell morphological profiling can be further refined to address its current limitations and unlock its full potential for studying complex cellular processes. Here, we presented cell death as a use case to show the utility of single-cell morphological profiling and informing drug discovery pipelines.

CRediT authorship contribution statement

Benjamin M. Frey: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **David Holmberg:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **Petter Byström:** Writing – review & editing, Validation, Investigation, Conceptualization. **Ebba Bergman:** Writing – review & editing, Methodology. **Polina Georgiev:** Writing – review & editing, Validation, Data curation. **Martin M. Johansson:** Writing – review & editing, Data curation. **Patrick Hennig:** Writing – review & editing, Validation, Methodology. **Jonne Rietdijk:** Writing – review & editing, Supervision, Data curation, Conceptualization. **Dan Rosén:** Writing – review & editing, Supervision, Conceptualization. **Jordi Carreras-Puigvert:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Ola Spjuth:** Writing – review & editing, Writing – original draft, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Funding

O.S. acknowledges funding from the Swedish Research Council (Grants 2020-03731, 2020-01865, 2024-04576, and 2024-03566), FORMAS (Grants 2022-00940 and 2025-02678), Swedish Cancer Foundation (22 2412 Pj 03 H), Horizon Europe (Grant Agreements 101057014 (PARC) and 101057442 (REMEDI4ALL)), and the Swedish strategic research programme eSSENCE.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Compounds were provided by Beactica Therapeutics. The computations were enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS), partially funded by the Swedish Research Council through grant agreement no. 2022-06725, resources provided by Uppsala University at UPPMAX, and the Berzelius resource provided by the Knut and Alice Wallenberg Foundation at the National Supercomputer Centre. The Authors acknowledge the support from the Chemical Biology Consortium Sweden (CBCS), the node at Uppsala University. CBCS is a national research infrastructure funded by the Swedish Research Council (dr.nr.2021-00179) and SciLifeLab.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ailsci.2026.100181.

Data and code availability

The code to run the analysis and generate the figures can be found on Github (https://github.com/pharmbio/sc_celldeath). The feature profiles, metrics, and splits to run the analysis and supervised training can be found on FigShare: https://doi.org/10.17044/scilifelab.28202864.

References

[1] Lotfollahi Mohammad, Klimovskaia Susmelj Anna, De Donno Carlo, Hetzel Leon, Ji Yuge, Ibarra Ignacio L, Srivatsan Sanjay R, Naghipourfar Mohsen, Daza Riza M, Martin Beth, et al. Predicting cellular responses to complex perturbations in high-throughput screens. Mol Syst Biol. 2023;19(6):e11517.

[2] Bunne Charlotte, Stark Stefan G, Gut Gabriele, del Castillo Jacobo Sarabia, Levesque Mitch, Lehmann Kjong-Van, Pelkmans Lucas, Krause Andreas, Rätsch Gunnar. Learning single-cell perturbation responses using neural optimal transport. Nature Methods 2023;20(11):1759–68.

[3] Peidli Stefan, Green Tessa D, Shen Ciyue, Gross Torsten, Min Joseph, Garda Samuele, Yuan Bo, Schumacher Linus J, Taylor-King Jake P, Marks Debora S, et al. scPerturb: harmonized single-cell perturbation data. Nature Methods 2024;1–10.

[4] Wu Yueshu, Yang Jun, Xu Gang, Chen Xiaolin, Qu Xiaochen. Integrated analysis of single-cell and bulk RNA sequencing data reveals prognostic characteristics of lysosome-dependent cell death-related genes in osteosarcoma. BMC Genomics 2024;25(1):379.

[5] Yao Wenchao, Liu Xuxu, He Yuanhang, Tian Maolan, Lu Shixin, Wang Qiang, Zheng Yi, Lv Zhenyi, Hao Chenjun, Xue Dongbo, Meng Xianzhi. ScRNA-seq and bulk RNA-seq reveal the characteristics of ferroptosis and establish a risk signature in cholangiocarcinoma. Mol Ther - Oncolytics 2022;27:48–60.

[6] Olofsson Karl, Carannante Valentina, Takai Madoka, Önfelt Björn, Wiklund Martin. Single cell organization and cell cycle characterization of DNA stained multicellular tumor spheroids. Sci Rep 2021;11(1):17076.

[7] Carpenter Anne E, Jones Thouis R, Lamprecht Michael R, Clarke Colin, Kang In Han, Friman Ola, Guertin David A, Chang Joo Han, Lindquist Robert A, Moffat Jason, Golland Polina, Sabatini David M. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biology 2006;7(10):R100.

[8] Haslum Johan Fredin, Lardeau Charles, Karlsson Johan, Turkki Riku, Leuchowius Karl-Johan, Smith Kevin, Müllers Erik. Cell Painting-based bioactivity prediction boosts high-throughput screening hit-rates and compound diversity. 2023, http://dx.doi.org/10.1101/2023.04.03.535328, bioRxiv.

[9] Rietdijk Jonne, Tampere Marianna, Pettke Aleksandra, Georgiev Polina, Lapins Maris, Warpman-Berglund Ulrika, Spjuth Ola, Puumalainen Marjo-Riitta, Carreras-Puigvert Jordi. A phenomics approach for antiviral drug discovery. BMC Biol 2021;19(1):156.

[10] Harrison Philip John, Gupta Ankit, Rietdijk Jonne, Wieslander Håkan, Carreras-Puigvert Jordi, Georgiev Polina, Wählby Carolina, Spjuth Ola, Sintorn Ida-Maria. Evaluating the utility of brightfield image data for mechanism of action prediction. PLoS Comput Biol 2023;19(7). e1011323.

[11] Stossi Fabio, Singh Pankaj K, Marini Michela, Safari Kazem, Szafran Adam T, Rivera Tostado Alejandra, Candler Christopher D, Mancini Maureen G, Mosa Elina A, Bolt Michael J, Labate Demetrio, Mancini Michael A. SPACe: an open-source, single-cell analysis of Cell Painting data. Nat Commun 2024;15(1):10170.

[12] Kraus Oren, Kenyon-Dean Kian, Saberian Saber, Fallah Maryam, McLean Peter, Leung Jess, Sharma Vasudev, Khan Ayla, Balakrishnan Jia, Celik Safiye, et al. Masked Autoencoders for Microscopy are Scalable Learners of Cellular Biology. 2024, http://dx.doi.org/10.48550/arXiv.2404.10242, arXiv.

[13] Kim Vladislav, Adaloglou Nikolaos, Osterland Marc, Morelli Flavio M, Halawa Marah, K´onig Tim, Gnutt David, Zapata Paula A Marin. Self-supervision advances morphological profiling by unlocking powerful image representations. 2024, http://dx.doi.org/10.1101/2023.04.28.538691, bioRxiv.

[14] Tian Guangyan, Harrison Philip J, Sreenivasan Akshai P, Carreras-Puigvert Jordi, Spjuth Ola. Combining molecular and cell painting image data for mechanism of action prediction. Artif Intell the Life Sci 2023;3:100060.

[15] Moshkov Nikita, Bornholdt Michael, Benoit Santiago, Smith Matthew, McQuin Claire, Goodman Allen, Senft Rebecca A, Han Yu, Babadi Mehrtash, Horvath Peter, Cimini Beth A, Carpenter Anne E, Singh Shantanu, Caicedo Juan C. Learning representations for image-based profiling of perturbations. Nat Commun 2024;15(1):1594.

[16] Caron Mathilde, Touvron Hugo, Misra Ishan, Jégou Hervé, Mairal Julien, Bojanowski Piotr, Joulin Armand. Emerging Properties in Self-Supervised Vision Transformers. 2021, http://dx.doi.org/10.48550/arXiv.2104.14294, arXiv.

[17] He Kaiming, Chen Xinlei, Xie Saining, Li Yanghao, Dollár Piotr, Girshick Ross. Masked Autoencoders Are Scalable Vision Learners. 2021, http://dx.doi.org/10.48550/arXiv.2111.06377, arXiv.

[18] Pfaendler Ramon, Hanimann Jacob, Lee Sohyon, Snijder Berend. Self-supervised vision transformers accurately decode cellular state heterogeneity. 2023, http://dx.doi.org/10.1101/2023.01.16.524226, bioRxiv.

[19] Doron Michael, Moutakanni Théo, Chen Zitong S, Moshkov Nikita, Caron Mathilde, Touvron Hugo, Bojanowski Piotr, Pernice Wolfgang M, Caicedo Juan C. Unbiased single-cell morphology with self-supervised vision transformers. 2023, http://dx.doi.org/10.1101/2023.06.16.545359, bioRxiv.

[20] Gupta Ankit, Wefers Zoe, Kahnert Konstantin, Hansen Jan N, Leineweber Will, Cesnik Anthony, Lu Dan, Axelsson Ulrika, Navarro Frederic Ballllosera, Karaletsos Theofanis, Lundberg Emma. SubCell: Vision foundation models for microscopy capture single-cell biology. 2024, http://dx.doi.org/10.1101/2024.12.06.627299, bioRxiv.

[21] Caicedo Juan C, Cooper Sam, Heigwer Florian, Warchal Scott, Qiu Peng, Molnar Csaba, Vasilevich Aliaksei S, Barry Joseph D, Bansal Harmanjit Singh, Kraus Oren, Wawer Mathias, Paavolainen Lassi, Herrmann Markus D, Rohban Mohammad, Hung Jane, Hennig Holger, Concannon John, Smith Ian, Clemons Paul A, Singh Shantanu, Rees Paul, Horvath Peter, Linington Roger G, Carpenter Anne E. Data-analysis strategies for image-based cell profiling. Nature Methods 2017;14(9):849–63.

[22] Pahl Axel, Sch´olermann Beate, Lampe Philipp, Rusch Marion, Dow Mark, Hedberg Christian, Nelson Adam, Sievers Sonja, Waldmann Herbert, Ziegler Slava. Morphological subprofile analysis for bioactivity annotation of small molecules. Cell Chem Biol. 2023.

[23] Lippincott Michael J, Tomkinson Jenna, Bunten Dave, Mohammadi Milad, Kastl Johanna, Knop Johannes, Schwandner Ralf, Huang Jiamin, Ongo Grant, Robichaud Nathaniel, et al. A morphology and secretome map of pyroptosis. Mol Biol Cell 2025. mbc.E25–03–0119.

[24] Chen Daniel, Plott Tempest, Wiest Michael, Van Trump Will, Komalo Ben, Nguyen Dat, Marsh Charlie, Heinrich Jarred, Fuller Colin J, Nicolaisen Lauren, Cambronero Elisa, Nguyen An, Elabd Christian, Rubbo Francesco, DeVay Jacobson Rachel. A combined AI and cell biology approach surfaces targets and mechanistically distinct Inflammasome inhibitors. IScience 2024;27(12):111404.

[25] Schorpp Kenji, Bessadok Alaa, Biibosunov Aidin, Rothenaigner Ina, Strasser Stefanie, Peng Tingying, Hadian Kamyar. CellDeathPred: a deep learning framework for ferroptosis and apoptosis prediction based on cell painting. Cell Death Discov 2023;9(1):1–11.

[26] Stirling David R, Swain-Bowden Madison J, Lucas Alice M, Carpenter Anne E, Cimini Beth A, Goodman Allen. CellProfiler 4: improvements in speed, utility and usability. BMC Bioinformatics 2021;22(1):433.

[27] Galluzzi Lorenzo, Vitale Ilio, Aaronson Stuart A, Abrams John M, et al. Molecular mechanisms of cell death: recommendations of the nomenclature committee on cell death 2018. Cell Death & Differ 2018;25(3):486–541. http://dx.doi.org/10.1038/s41418-017-0012-4.

[28] Kerr JFR, Wyllie AH, Currie AR. Apoptosis: A Basic Biological Phenomenon with Wideranging Implications in Tissue Kinetics. Br J Cancer 1972;26(4):239–57.

[29] Conrad Marcus, Angeli José Pedro Friedmann, Vandenabeele Peter, Stockwell Brent R. Regulated necrosis: disease relevance and therapeutic opportunities. Nat Rev Drug Discov 2016;15(5):348–66.

[30] Stockwell Brent R. Ferroptosis turns 10: Emerging mechanisms, physiological functions, and therapeutic applications. Cell 2022;185(14):2401–21.

[31] Yu Pian, Zhang Xu, Liu Nian, Tang Ling, Peng Cong, Chen Xiang. Pyroptosis: mechanisms and diseases. Signal Transduct Target Ther 2021;6(1):1–21.

[32] Shen Sikou, Shao Yina, Li Chenghua. Different types of cell death and their shift in shaping disease. Cell Death Discov 2023;9(1):1–12.

[33] Tesniere A, Schlemmer F, Boige V, Kepp O, Martins I, Ghiringhelli F, Aymeric L, Michaud M, Apetoh L, Barault L, Mendiboure J, Pignon J-P, Jooste V, van Endert P, Ducreux M, Zitvogel L, Piard F, Kroemer G. Immunogenic death of colon cancer cells treated with oxaliplatin. Oncogene 2010;29(4):482–91.

[34] Kroemer Guido, Galluzzi Lorenzo, Kepp Oliver, Zitvogel Laurence. Immunogenic Cell Death in Cancer Therapy. Annu Rev Immunol 2013;31(Volume 31, 2013):51–72.

[35] Willis Clinton, Nyffeler Johanna, Harrill Joshua. Phenotypic Profiling of Reference Chemicals across Biologically Diverse Cell Types Using the Cell Painting Assay. SLAS DISCOVERY: Adv Sci Drug Discov 2020;25(7):755–69.

[36] Francisco Rodríguez María Andreína, Carreras Puigvert Jordi, Spjuth Ola. Designing microplate layouts using artificial intelligence. Artif Intell the Life Sci 2023;3:100073.

[37] Bray Mark-Anthony, Singh Shantanu, Han Han, Davis Chadwick T, Borgeson Blake, Hartland Cathy, Kost-Alimova Maria, Gustafsdottir Sigrun M, Gibson Christopher C, Carpenter Anne E. Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. Nat Protoc 2016;11(9):1757–74.

[38] Stringer Carsen, Wang Tim, Michaelos Michalis, Pachitariu Marius. Cellpose: a generalist algorithm for cellular segmentation. Nature Methods 2021;18(1):100–6.

[39] Singh S, Bray M-A, Jones Tr, Carpenter Ae. Pipeline for illumination correction of images for high-throughput microscopy. J Microsc 2014;256(3):231–6.

[40] Gregor Way. Pycytominer: Data processing functions for profiling perturbations. 2022.

[41] Kalinin Alexandr A, Arevalo John, Serrano Erik, Vulliard Loan, Tsang Hillary, Bornholdt Michael, Muñoz Alán F, Sivagurunathan Suganya, Rajwa Bartek, Carpenter Anne E, Way Gregory P, Singh Shantanu. A versatile information retrieval framework for evaluating profile strength and similarity. 2025, http://dx.doi.org/10.1101/2024.04.01.587631, bioRxiv.

[42] Wolf F Alexander, Angerer Philipp, Theis Fabian J. SCANPY: large-scale single-cell gene expression data analysis. Genome Biology 2018;19(1):15.

[43] Kotyan Shashank, Ueda Tatsuya, Vargas Danilo Vasconcellos. K* Distribution: Evaluating the Latent Space of Deep Neural Networks using Local Neighborhood Analysis. 2024, http://dx.doi.org/10.48550/arXiv.2312.04024, arXiv.

[44] Erickson Nick, Mueller Jonas, Shirkov Alexander, Zhang Hang, Larroy Pedro, Li Mu, Smola Alexander. AutoGluon-Tabular: Robust and Accurate AutoML for Structured Data. 2020, http://dx.doi.org/10.48550/arXiv.2003.06505, arXiv.

[45] Chen Tianqi, Guestrin Carlos. XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. 2016, p. 785–94.

[46] Dorogush Anna Veronika, Ershov Vasily, Gulin Andrey. CatBoost: gradient boosting with categorical features support. 2018, http://dx.doi.org/10.48550/arXiv.1810.11363, arXiv.

[47] Breiman Leo. Random Forests. Mach Learn 2001;45(1):5–32.

[48] Korsunsky Ilya, Millard Nghia, Fan Jean, Slowikowski Kamil, Zhang Fan, Wei Kevin, Baglaenko Yuriy, Brenner Michael, Loh Po-ru, Raychaudhuri Soumya. Fast, sensitive and accurate integration of single-cell data with Harmony. Nature Methods 2019;16(12):1289–96.

[49] Mueller Sandra, Schittenhelm Marcus, Honecker Friedemann, Malenke Elke, Lauber Kirsten, Wesselborg Sebastian, Hartmann Joerg T, Bokemeyer Carsten, Mayer Frank. Cell-cycle progression and response of germ cell tumors to cisplatin in vitro. Int J Oncol 2006;29(2):471–9.

[50] Matuszewski Damian J, W´ahlby Carolina, Puigvert Jordi Carreras, Sintorn Ida-Maria. PopulationProfiler: A Tool for Population Analysis and Visualization of Image-Based Cell Screening Data. PLoS One 2016;11(3):e0151554.

[51] Roukos Vassilis, Pegoraro Gianluca, Voss Ty C, Misteli Tom. Cell cycle staging of individual cells by fluorescence microscopy. Nat Protoc 2015;10(2):334–48.

[52] Dehghan Sepehr, Kheshtchin Nasim, Hassannezhad Shaghayegh, Soleimani Maryam. Cell death classification: A new insight based on molecular mechanisms. Exp Cell Res 2023;433(2):113860.

[53] Chen Yao, Li Xiaohua, Yang Minfeng, Liu Song-Bai. Research progress on morphology and mechanism of programmed cell death. Cell Death Dis 2024;15(5):1–13.

[54] Tanevski Jovan, Flores Ricardo Omar Ramirez, Gabor Attila, Schapiro Denis, Saez-Rodriguez Julio. Explainable multiview framework for dissecting spatial relationships from highly multiplexed data. Genome Biology 2022;23(1):97.

[55] Selvaraju Ramprasaath R, Cogswell Michael, Das Abhishek, Vedantam Ramakrishna, Parikh Devi, Batra Dhruv. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. 2019, http://dx.doi.org/10.48550/arXiv.1610.02391, arXiv.

[56] Dong Mingze, Wang Bao, Wei Jessica, de O. Fonseca Antonio H, Perry Curtis J, Frey Alexander, Ouerghi Feriel, Foxman Ellen F, Ishizuka Jeffrey J, Dhodapkar Rahul M, van Dijk David. Causal identification of single-cell experimental perturbation effects with CINEMA-OT. Nature Methods 2023;20(11):1769–79.

[57] Baran Yael, Bercovich Akhiad, Sebe-Pedros Arnau, Lubling Yaniv, Giladi Amir, Chomsky Elad, Meir Zohar, Hoichman Michael, Lifshitz Aviezer, Tanay Amos. MetaCell: analysis of single-cell RNA-seq data using K-nn graph partitions. Genome Biol. 2019;20(1):206.

[58] Dijk Robert van, Arevalo John, Babadi Mehrtash, Carpenter Anne E, Singh Shantanu. Capturing cell heterogeneity in representations of cell populations for image-based profiling using contrastive learning. 2023, http://dx.doi.org/10.1101/2023.11.14.567038, bioRxiv.

[59] Dosovitskiy Alexey, Beyer Lucas, Kolesnikov Alexander, Weissenborn Dirk, Zhai Xiaohua, Unterthiner Thomas, Dehghani Mostafa, Minderer Matthias, Heigold Georg, Gelly Sylvain, Uszkoreit Jakob, Houlsby Neil. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. 2021, http://dx.doi.org/10.48550/arXiv.2010.11929, arXiv.

[60] Althnian Alhanoof, AlSaeed Duaa, Al-Baity Heyam, Samha Amani, Dris Alanoud Bin, Alzakari Najla, Abou Elwafa Afnan, Kurdi Heba. Impact of Dataset Size on Classification Performance: An Empirical Evaluation in the Medical Domain. Appl Sci 2021;11(2):796.

[61] Fradkin Philip, Azadi Puria, Suri Karush, Wenkel Frederik, Bashashati Ali, Sypetkowski Maciej, Beaini Dominique. How Molecules Impact Cells: Unlocking Contrastive PhenoMolecular Retrieval. 2024, http://dx.doi.org/10.48550/arXiv.2409.08302, arXiv.

[62] Sanchez-Fernandez Ana, Rumetshofer Elisabeth, Hochreiter Sepp, Klambauer G$^{\prime}$unter. CLOOME: contrastive learning unlocks bioimaging databases for queries with chemical structures. Nat Commun 2023;14(1):7339.