Label-Free Live-Cell Imaging improves Mode of Action Classification

Edvin Forsgren[1,*], Jonne Rietdijk[2], David Holmberg[2], Martin M. Johansson[2], Jordi Carreras-Puigvert[2], Johan Trygg[1,3], Gillian Lovell[4], Ola Spjuth[2], and Pär Jonsson[3]

[1]Computational Life Science Cluster (CLiC), Department of Chemistry, Umeå University, Umeå, Sweden
[2]Department of Pharmaceutical Biosciences and Science for Life Laboratory, Uppsala University, Uppsala, Sweden
[3]Sartorius Corporate Research, Umeå, Sweden
[4]Sartorius BioAnalytics, Royston, UK
[*]Correspondence: edvin.forsgren@umu.se

SUMMARY

Morphological profiling is a powerful method for identifying the modes of action (MOAs) of chemical compounds. However, most approaches rely on fixed-cell assays that capture only a single time point, missing the dynamic nature of cellular responses. While live-cell imaging captures these temporal effects, its potential for MOA classification using high-dimensional features remains unexplored.

We strategically convert large-scale time-series image data (>82,000 images) into interpretable phenotypic trajectories and show that label-free live-cell imaging captures meaningful temporal phenotypic signatures that improve MOA classification compared to single-time-point analyses. The results are consistent across two cell lines and six MOA groups.

This workflow enables mechanistic insight from live-cell images and supports the growing interest in label-free methods, which offer advantages in preparation time and cost, and preserve the native state of cells. Our findings demonstrate that live-cell imaging, combined with data-driven analysis, offers a powerful tool for high-throughput drug screening.

KEYWORDS

Live-Cell Imaging, Label-Free Imaging, MOA Classification, Morphological Profiling, Deep Learning, Time Series Analysis

1 Introduction

Time-series data capture the essential temporal dynamics of biological systems, yet extracting meaningful insights from high-dimensional, time-lapse cellular images remains a significant challenge in modern drug discovery[1,2]. The ability to analyse how cells respond to compounds over time, rather than at arbitrary endpoints, represents a critical advancement for understanding therapeutic modes of action (MOA) and predicting clinical outcomes[3].

One powerful tool for studying the MOA of drugs involves capturing and analysing cellular phenotypic changes in response to treatment - an approach known as image-based morphological profiling[4,5]. By leveraging high-content imaging techniques such as the Cell Painting assay[6–8], researchers can generate rich, unbiased datasets that reveal how different compounds alter cellular structures and organelle organization[7]. These morphological fingerprints can be compared across a reference database of known drug effects, enabling the identification of shared phenotypic signatures to infer MOA and support virtual screening[9].

Today, morphological profiling is typically performed as a single end point assay, where compounds are applied for a set period, such as 48 hours, followed by cell fixation and staining using fluorescent dyes or antibodies. This requires complex staining protocols, potentially introducing technical variations due to factors such as reagent lots, processing times, equipment calibration, and experimental platforms, leading to batch effects between plates and laboratories. Additionally, an endpoint assay only provides a static snapshot of cellular states, potentially missing important effects that occurred earlier or have not yet happened and thus missing important time-dependent dynamics.

Biological processes are inherently dynamic. Cells continuously respond to their environment, activate complex signalling cascades, undergo division, differentiate, and ultimately die. These processes occur across varying timescales, from seconds for ion channel activity to hours for gene expression changes to days for phenotypic transformations[10]. Capturing this temporal complexity is fundamental to understanding both normal physiology and disease states. Living cells represent dynamic, adaptive systems rather than static entities, making time-resolved analysis essential for accurate biological insights.

In cancer drug development, temporal monitoring of drug-treated cells provides insights into whether cells relapse, undergo senescence, or develop alternative survival pathways, which are key factors for designing long-lasting cancer treatments[3]. Live-cell imaging has revealed new insights into tumour and T cell behaviour, leading to the discovery of novel druggable targets to optimize engineered T cell function for improved tumour targeting[11], and identified effective drug combinations for treating heterogeneous tumour populations[12,13].

To capture temporal dynamics in live cells, researchers commonly use permeable, live-cell-compatible stains. However, these stains can become unstable over time, and repeated imaging exposures may lead to phototoxicity due to the damaging fluorescent light. To minimize these effects, assay conditions must be carefully optimized, leading to a trade-off between assay duration, signal quality, and cell health[14].

To circumvent challenges associated with stain instability and phototoxicity, whilst also reducing cost and labour, label-free imaging offers a compelling alternative[15]. It preserves native cellular states and avoids artefacts introduced by labelling. When combined with AI-driven segmentation and feature extraction, label-free imaging provides a powerful, non-invasive approach for live-cell analysis. While label-free live-cell datasets exist[16], focus has been mostly on segmentation or cell tracking[17]. Many studies rely on a limited set of handcrafted features for characterizing cellular phenotypes, or on quantitative phase imaging (QPI), which lacks the fine subcellular detail captured by phase contrast microscopy[18,19]. Exploring cellular phenotypes with high-dimensional features on label-free live-cell time series data is not widely explored yet.

AI and deep learning is revolutionizing the possibilities of morphological profiling by enabling the extraction of complex, high-dimensional features from imaging data with unprecedented accuracy and scalability[4]. With these new technologies, high-dimensional features can be readily extracted from label-free cell imaging. And unlike traditional analysis methods that rely on predefined features, deep learning models can automatically learn intricate phenotypic patterns, improving sensitivity in detecting subtle cellular changes[20,21].

In this paper, we utilize deep learning to convert time series of label-free live-cell images into interpretable profiles showcasing compounds dynamic effects over time by introducing a multi-model workflow. The workflow consists of pre-trained deep learning models for cell segmentation, live/dead classification, and single cell feature extraction. Our results reveal that supervised time-dependent phenotypic modelling provides valuable insights into the temporal dynamics of compound effects and significantly enhances accuracy in MOA classification compared to a single time point analysis.

2 Materials and Methods

2.1 Cell Assay and Live-Cell Imaging

Prior to assay, compound plates were thawed for 30 minutes at room temperature in a sterile biological safety cabinet. Cells (A549-WT, U2-OS-WT; Supplementary Table 1) were cultured in T75 cell culture flasks and harvested on the day of assay by aspiration of media, washing with 10 mL PBS, and addition of 1.5 mL Trypsin-EDTA solution. Once detached, cells were suspended in fresh media and counted (Invitrogen Countess 3 FL). Cell suspensions were formed in 50 mL centrifugation tubes (2.5 E+04 cells/mL, 13 mL per plate). After mixing well, the cell suspension was added into the 384-well compound plate in 30 µL/well volume using a manual multichannel pipette. Plates were centrifuged for 1 minute at 100xg to align cells at the bottom of the well, then placed into an Incucyte® SX5 Live Cell Analysis System equipped with a Green/Orange/NIR optical module. Images were acquired using Adherent Cell-by-Cell acquisition at 20x magnification every 2h for 3 days, with a single image site per well. Images were analysed using AI Cell Health Analysis® which provides individual cell segmentations and classification as live or dead[22,23] (Supplementary Table 2).

2.2 Compound treatments

We selected compounds belonging to six diverse MOA classes that we believed would be reasonably separable and expected to be either fast-acting or slow-acting. Specifically, we included AKT inhibitor (AKT), CDK inhibitor (CDK), HDAC inhibitor (HDAC), MAP kinase inhibitor (MAPK), PARP inhibitor (PARP) and tubulin polymerization inhibitor (TUB). For each of these MOAs we selected twelve compounds from the SPECS repurposing library. To account for differences in compound potency and toxicity, we used internal data from compounds tested at 10 µM to ensure detectable effects without inducing toxicity (data not published). One compound from the MAPK group was excluded, see Supplementary Figure 1.

Each compound was tested in fifteen replicates, consisting of five technical replicates (i.e., compound treatments on the same plate) and three biological replicates (i.e., treatments on different plates). Each plate included fifteen wells containing only the solvent, DMSO (Dimethyl Sulfoxide), as negative controls.

The plate layout for the six microplates, three per cell line, was done using PLAID (Plate Layouts using Artificial Intelligence Design[24], a constraint-programming-based method designed to minimize unwanted well-position and plate effects. Compound handling was performed by the SciLifeLab Compound Center (CBCS, Solna, Stockholm). Briefly, chemicals were solubilized in either DMSO or water at stock concentrations between 2 and 10 mM, as specified in Supplementary Table 3. Compounds were dispensed in nanoliter volumes using the Echo liquid handler into 384-well PhenoPlates (PerkinElmer, cat. no. 6057302) and stored at -20℃ prior to experimentation.

2.3 Characterisation of live cell assay

The live-cell imaging data was used to examine the effect of these compounds on A549 and U2OS cells. Untreated (DMSO) A549 cells displayed high viability throughout the three-day time frame of the assay, and proliferation was observed as the confluence increased from $\sim$12% to over 75%. The single time point (48h) scatter plots, figure 2A, indicate that a small number of compounds strongly affected A549 viability, showing a cluster of compounds in the bottom left quadrant that decreased viability, inducing cell death. Another cluster of compounds can be

observed in which cells have reduced confluence ($<$25%) but maintain viability in the range of 60-80%, suggesting a cytostatic mechanism and the possibility of cell cycle arrest. The remaining compounds retain $>$80% viability in most cases with a range of effects on cell growth indicated by confluence values ranging from 25 to 80%. Figure 2E displays negative control A549 cells with a normal healthy adherent cell morphology, and then the effect of treatments showing one example from each MOA. Afuresertib (AKT) subtly altered the shape, with an increase in circularity and texture. Albendazole (TUB) generated ruffling in the plasma membrane boundary, with a higher proportion of visible mitotic cells. Aminopurvalanol (CDK) treated cells displayed an increase in contrast around the nuclear envelope. Binimetinib (MAPK) increased cell area with prominent focal adhesion points. M344 (HDAC) and Talazoparib (PARP) both increased cell area; with M344-treated cells appearing ruffled at the boundary while Talazoparib-treated cells maintained a more circular appearance.

Untreated (DMSO negative control) U2OS cells also displayed high viability throughout the time course of the assay, however, the rate of proliferation was lower than that of A549, reaching just under 50% by the end of three days as displayed in figure 2D. In contrast to A549 cells, the single time point (48h) scatter plot, figure 2C, indicates that fewer compounds affected U2OS viability. Figure 2E shows that untreated U2OS cells display a heterogeneous, adherent morphology with lower contrast relative to A549 cells. Treatment with Afursertib (AKT) induced a very subtle change in morphology with a slight increase in cytosolic texture which can indicate stress. Albendazole (TUB) treatment increased the proportion of cells in mitosis with remaining cells appearing more rounded. Cells treated with Aminopurvalanol (CDK) appeared enlarged, flattened, and rounded while those treated with Binimetinib (MAPK) and Talazoparib (PARP) experienced increased area and roundness. HDAC inhibitor M344 (HDAC) resulted in more cells appearing elongated and with higher contrast.

In figure 2F and G, scatter plots display unsupervised UMAP[25] embeddings (n_neighbours=25, min_dist=0.1, metric="euclidean") of cell morphology features extracted using DINOv2-ViT-B14 (See Section 3.1) from wells containing A549 and U2OS cells, coloured by MOA group. The data span 24h to 72h, with each point representing a well at a specific time point. For A549, figure 2F, distinct clusters of MOA groups emerge, though many MOAs remain heterogeneous, forming scattered clusters, for example, AKT (green), CDK (blue), and MAPK (yellow). In contrast U2OS, figure 2G, show less distinct clustering overall, with only a few recognizable groups such as TUB (purple), clusters of AKT, and HDAC (red). UMAP visualizations provide an overview of the data, and colouring points by known metadata helps reveal patterns, both desired (MOA/compound clusters) and undesired (batch effects). However, while these embeddings highlight structure, interpreting the specific effects of different MOA groups on cells remains challenging since unsupervised methods such as UMAPs are designed to capture overall variance of the dataset and not specific differences between groups.

2.4 Model and Training Details

The ANNs used in this work are fully connected deep learning architectures implemented to predict soft MOA labels based on two different feature sets. One set consisting 768 deep learning features (explained in more detail in section 3) and one set with twelve cell properties such as area, confluence, eccentricity, cell count, and viability. For the deep learning features, the model takes the vector of 768 features as input followed by two hidden layers, consisting of 512 and 256 neurons, and an output layer with six neurons, one for each MOA in our dataset. For the cell properties, the network takes 12 features as input and two hidden layers with 8 neurons each, and with the same output layer of six MOA-specific neuron. The networks use LeakyReLU activation functions in the hidden layers, and a high dropout rate of 0.5 to reduce overfitting. The

output layer has no activation function and thus outputs a soft score for each of the six MOAs. The model uses L1 Loss and the AdamW optimizer with a learning rate of 0.0001. The models were implemented in pytorch[26] and were trained for 200 epochs with no early stopping. For our time-dependent modelling approach with DINO features, the full CV scheme takes $\sim$20 minutes to run with parallelization on one NVIDIA GeForce RTX 3090 and a batch size of 16.

3 Data Analytical Workflow

In this study, we introduce a time-dependent modelling framework for analysing time series data, referred to as the "time series approach." This method utilizes two distinct feature sets: deep learning-based features, termed "DINO features," and a set of "cell properties" encompassing metrics such as area, shape, and cell count. The inclusion of cell properties serves to assess whether DINO features capture subtle morphological variations that standard cell properties might overlook. Additionally, we contrast the time series approach with a single time point analysis for both feature sets to examine the advantages of incorporating temporal data for MOA classification and temporal insights.

3.1 Pre-trained Models: Segmentation and Feature Extraction

Single-cell segmentation masks, derived from the label-free AI Cell Health Analysis[22,23], were employed to extract individual cells. These cells were cropped and zero-padded to a resolution of 224×224 pixels. Cells exceeding this size or located at image boundaries were excluded, leading to the removal of approximately 20% of the total cells. A Chi-squared test ($p = 0.997$) confirmed no significant variation in exclusion proportions across the six MOA groups, supporting the null hypothesis H_0 of MOA-independent exclusion.

The processed cell crops were then passed through a pre-trained DINOv2-ViT-B14 model, extracting 768 features per single cell[27–29]. Utilizing cell health labels from the AI Cell Health analysis, median well-aggregated features were computed for each time point, incorporating only viable cells. This provided a well-level representation per time point, exclusively reflecting the morphology of living cells. Each plate was then normalized using a z-score calculated from negative controls (DMSOs) present on the same plate. These processed outputs constitute the DINO feature set. Additionally, a well-level feature set capturing cell properties such as size, area, and health classification was extracted from the AI Cell Health Analysis software.

3.2 Data-Driven Models: Artificial Neural Networks (ANNs)

Based on these median well features, we trained two ANNs: one for capturing early MOA-specific phenotypes (6-14 hours) and another for late phenotypes (46-54 hours). Each ANN was trained using data within its designated time window, along with negative control data from both windows. These models utilize median well features (DINO features or cell properties) as inputs, with each sample representing a specific well and time point. The ANNs have six outputs, each corresponding to a distinct MOA group. During training, wells exposed to compounds associated with a given MOA were assigned an output value of one for the corresponding MOA, while the remaining outputs were set to zero.

The models employ a soft classifier approach, which assigns values on a continuous scale rather than discrete class labels. A value of zero indicates non-membership in an MOA group, while one represents membership. This approach introduces flexibility, allowing predictions to

fall below zero or above one, ensuring that compounds with weaker but similar effects are not strictly classified as belonging to a group but rather assigned intermediate values.

3.3 Eq. Scores and Eq. Profiles

Once trained, the ANN models were applied to unseen samples, generating predictions for each MOA group across all wells and time points. These predictions are referred to as Equivalence Scores (Eq. scores)[30], with one Eq. score assigned per MOA group. Eq. scores from all time points for a given well form Eq. profiles, which serve as temporal signatures for MOA classification (see Figure 3).

3.4 Smoothing with Exponentially Weighted Moving Average

To leverage temporal information while mitigating random fluctuations, Eq. profiles were smoothed using Exponentially Weighted Moving Average (EWMA)[31]. This smoothing technique enhances the robustness of MOA classification by reducing noise and emphasizing true temporal trends.

3.5 MOA Classification: Cross-Validation Procedure

For model evaluation, a double-blind cross-validation (CV) procedure was implemented to simulate real-world conditions and prevent data leakage. The CV process was structured as follows:

- In each CV round, training and test sets were created. The training set comprised data from two plates, excluding wells treated with a set of six compounds (one per MOA group).
- The test set included all wells treated with these six compounds on the third plate, which was omitted from training.
- Each plate underwent 12 CV rounds, corresponding to the 12 available compounds per MOA.
- Given three plates per cell line, this resulted in a total of 36 CV rounds per cell line.

For the time series approach, where separate early and late models were trained, this led to a total of 72 models being trained per cell line. During each CV round, predictions for six MOA-specific Eq. scores were generated across all time points (0-72 hours) in the unseen test set. This resulted in two Eq. scores per MOA per well and time point—one from the early model and one from the late model. These Eq. scores were then smoothed using EWMA, producing refined Eq. profiles that capture MOA-related temporal dynamics.

For the single time point analysis, the same CV scheme was employed, but only a single model per round was trained, resulting in 36 models per cell line. Predictions were generated solely at the 48-hour time point, a commonly used endpoint in assays. Both DINO features and the cell property feature set were tested using the ANN architectures detailed in Section 2.4.

3.6 Detection Limits and MOA Assignment

To prevent the misclassification of compound-treated wells that do not significantly deviate from negative controls, detection limits were established. These thresholds ensure that compounds at suboptimal concentrations, which do not induce detectable morphological changes, are not falsely classified as exhibiting MOA-related effects.

Detection limits were determined from Eq. scores of negative control wells in the CV scheme (Figure 3). They were computed as the mean plus two standard deviations of the sum of the integrals of Eq. profiles (both early and late) for each MOA group, referred to as Eq. sum. A similar approach was applied to the single time point analysis, where detection limits were based on the average plus two standard deviations of MOA-specific Eq. scores for negative controls.

Once detection limits were established, the Eq. sum for each compound-treated well was computed by integrating and summing the early and late Eq. profiles. If the Eq. sum exceeded the detection threshold, the MOA group with the highest Eq. sum was assigned as the predicted label. This filtering mechanism ensured that only significantly altered wells were classified, reducing false positive rates. Notably, detection limits varied across MOA groups, reflecting differences in morphological response intensities.

3.7 Comparative Analysis: Time Series vs. Single Time Point Approach

To evaluate the effectiveness of the time series approach, we compared its performance against the single time point analysis for both feature sets. The same CV procedure and ANN architectures were used for training across both methods, with architecture modifications tailored for the DINO features and cell property feature sets.

By integrating Eq. profiles and detection limits, our workflow enables robust MOA classification while accounting for temporal dynamics, providing a more comprehensive understanding of cell responses compared to traditional single time point analyses.

4 Results

Here, we present a new approach that leverages time series live-cell imaging to generate interpretable profiles that capture the dynamic effects of compounds over time. We found that examining the full time series data of cell treatments results in a significantly higher accuracy for MOA classification compared to analysing a single time point.

4.1 Temporal Insights

Because each compound has a unique dose-response relationship, its effective concentration range varies. When screening thousands of compounds at a single concentration, there is no guarantee that every compound will be tested at a level where its effects are detectable at the same single time point. However, by capturing the full time series, early rapid responses to a specific MOA can still be identified, even if the cells ultimately undergo death. This allows for correlation between fast-acting, high-efficacy compounds and slower responses from compounds with the same MOA but lower potency by comparing different time points.

In addition to concentration dependency, compounds with certain MOAs may induce a transient or non-linear effect on the cells while others have a rapid and linear response. Such a response is seen for the AKT group which displays a rapid response which is sustained throughout the remainder of the time series, resulting in similar Eq. profiles for both early and late model predictions (figure 4). Noticeably for CCT128930 (AKT), the live-cell images from 8h and 48h share similar morphologies although the confluence has increased at the later time point, confirming what can be seen in the Eq. profiles for both cell lines. This is consistent with previously observed responses to AKT inhibitors, where the AKT signalling pathway is activated within minutes of inhibitor exposure[32].

A different example are the HDAC inhibitors that for A549 cells exhibit a low early Eq. profile but with a clear late profile. Because HDAC inhibitors are believed to regulate transcription through epigenetic modifications, which typically take time to manifest[33], it is expected to see such delayed effects.. For U2OS cells, the profiles show both clear early and late Eq. profiles showcasing a clear difference between the two cell lines.

Similarly to HDAC inhibitors, PARP inhibitors, which prevent the repair of single-strand DNA breaks[34], have only a very late onset and show the lowest classification accuracy (figures 5, 6, and 7). We hypothesize that their effects become apparent only after mutations accumulate over time, possibly explaining the delayed effects observed and might even be more evident at later time points than acquired in this dataset.

An example of a group that has clear, distinct early and late effects is the tubulin-targeting group (TUB compounds), which targets the cytoskeleton[35]. During mitosis, the cytoskeleton dynamically rearranges and this process is disrupted by treatment with these compounds resulting in an increase in mitotic index. This is observed at 8h in the image in figure 4. This stage is followed by the cells presenting ruffled plasma membrane boundaries and an altered morphology distinct from untreated cells (see figure 4). These two distinct morphologies are clearly seen in the Eq. profiles of Nocadazole (TUB) in figure 4 for both cell lines with clear early and late profiles. In addition, small, rounded, mitotic cells are frequent in live-cell images taken at 8h but at a later time point, 48h, the cells exhibit a ruffled cell plasma membrane instead.

To investigate the temporal effects of different compounds, we generated CV Eq. profiles based on Eq. scores, smoothing, and a CV scheme. These profiles, grouped by MOA, are shown in figure 5 for A549 and figure 6 for U2OS.

From the Eq. profiles we can see that the AKT compounds exhibit no clear early-late distinction, instead showing high Eq. scores for both consistently throughout the majority of the time course. The major exceptions for AKT are Buparsilib, and Triciribine. Where both display low AKT Eq. scores in both cell lines (figure 5 and 6). CDK compounds are overall more heterogeneous, yet like AKT, they lack a clear early-late distinction. What stands out are Abemaciclib and BMS-265246, demonstrating a notable early onset of AKT-late Eq. scores from hours 0 to 30.

Since the AKT and CDK pathways are both directly and indirectly connected, it has been shown that compounds inhibiting AKT (upstream) may indirectly reduce the activation of CDK 2, 4, and 6 (downstream)[36]. Additionally, the AKT-inhibitor AR-12 shows slight CDK-late Eq. scores in A549 cells. CDK4/6 inhibition has been previously reported to suppress AKT-mTOR signalling which might explain these findings[37].

Palbocicilib and Roscovitine exhibit little to no effect. In U2OS cells, most CDK compounds have low scores, with only LY2857785, RGB-286638, and TG-02 showing visible profiles. For HDAC compounds, clear late Eq. profiles along with less distinct early profiles are observed in A549 cells, figure 5, with Droxinostat and Pyroxamide showing lower scores than the rest. For U2OS cells both early and late profiles are more clearly visible in figure 6, with Droxinostat and Pyroxamide showing low scores with the addition of UF010 and Ricolinostat, which were also on the lower end for A549 cells.Suberohydroxamic acid (SBHA) displays a clear AKT-early Eq. profile. MAPK compounds differ significantly between the two cell lines. While many compounds show both early and late profiles separated over time in A549, U2OS exhibits low Eq. scores overall. PARP compounds generally present very subtle profiles, if any. For those showing a PARP profile, only late profiles are visible, consistent across both cell lines. TUB compounds display a clear early-late distinction in their Eq. profiles. Parbendazole has a less distinct profile but remains clearly noticeable in both cell lines.

A supervised modelling-design like this reveals far more details about each compounds effect compared to an unsupervised such as UMAP (figure 2F and G). However, when training supervised models, it must be done with care to avoid over-optimistic results and faulty conclusions. Therefore, to validate our models and findings, we consistently predict new compounds

on new plates to mimic a real world drug screening scenario. This conservative CV scheme is only necessary for the last part of our workflow, the supervised, time-dependent modelling with small ANNs because the rest of the models are pre-trained and do not need to be validated.

4.2 Time series vs Single time point - MOA Classification

Using the Eq. sums calculated from the Eq. profiles above, as well as the Eq. scores from models trained only on a single time point (48h), and their respective detection limits, we predict MOA for the compound-treated wells for models trained with high-dimensional DINO features and less complex models trained with cell properties. For the MOA classification, each well is classified into one of seven classes (AKT, CDK, HDAC, MAPK, PARP, TUB, and DMSO). However, only wells belonging to the six MOA classes are classified and included in the accuracy calculations. Wells from the seventh class, DMSO i.e. negative controls, are not included, as this class would by default have an accuracy of $\sim$97.5% based on the definition of the detection limit. Importantly, if a well's highest Eq. sum falls below the detection limit, it is classified as DMSO.

Using the cell properties, we achieved accuracies of 14% and 26% for A549- and U2OS-cells using data from a single time point (48h). With time series, we saw an improvement for the A549 cells achieving 26% accuracy but for the U2OS cells the accuracy was again 26%.

Figure 7 shows the MOA classification accuracies from the DINO features for the single time point and time series for the A549 and U2OS cells. In figure 7A, the predictions for A549 cells at a single time point (48h) yield a total accuracy of 60%. While AKT, CDK, HDAC, and TUB are classified with relatively high accuracy. MAPK and more so PARP, often fail to pass the detection limit and thus classified as negative controls (DMSO), 42.2% of MAPK and 72.3% of PARP. Meaning, these were not significantly different from the negative controls.

For U2OS single time point classification (figure 7B), the accuracy is lower at 49%, with increased misclassification across multiple MOAs. Most notably for CDK where fewer instances passes the detection limit and 34.4% ends up in the negative control class (DMSO). MAPK and PARP are again hard to distinguish from negative controls leading to 22.2% and 50.6% classified as DMSO respectively.

Time series classification significantly improves accuracy for both cell lines across all MOAS with $p = 0.0012$ for A549 and $p = 0.0183$ for U2OS. In figure 7C, the accuracy for A549 increases to 79%, with strong improvements in classifying PARP, from 6.1% accuracy in A to 38.3% in C, and a pronounced reduction in MAPK misclassification into DMSO, from 38.3% correctly classified to 65.0%. Similarly, U2OS classification improves to 62%, with AKT accuracy increasing from 67.8% in B to 90.6% in D. Time series classification enhances overall performance, highlighting the advantage of incorporating temporal data for more accurate MOA classification, significantly improving accuracy for all MOAs for both cell lines.

The presented workflow is highly modular serving two main purposes. First, it allows for easy substitution of models in each step when improvements become available. Segmentation-, live/dead classifying-, and feature extraction models are continuously improved, and as better models enter the scene, our proposed workflow will still be applicable and with, most likely, improved results. Second, it acts as a means of "concentrating" relevant information. Meaning, only the parts of the data we want to focus on are passed onto the next step. The segmentation model identifies the cells within images and by removing the background, we make sure the feature extraction model focus solely on the cells. Then, by only including the living cells based on the live/dead classification model, we limit the impact of cell death and focus on identifying more subtle morphological changes. Thus, get more control over what goes into our median well profiles. Using these profiles, we train supervised, time-dependent models to create interpretable Eq. profiles that are both useful for temporal insights and MOA prediction with a

thorough double-blind CV scheme. Doing the same CV scheme with an image-based classifier, i.e. a model that takes raw cell images as input and outputs a MOA label, would be considerably more computationally intensive highlighting another benefit of our modular approach.

5 Conclusion

This paper introduces a novel workflow for analysing complex high-dimensional time series data from label-free live-cell imaging, yielding interpretable and biologically relevant results. The workflow includes multiple pre-trained AI models to segment cells, classify them as live or dead, and extract high-dimensional single-cell features. These features serve as input to custom-trained ANN models, whose output reveals temporal drug responses, forming the basis for MOA classification.

Comparing MOA classification accuracy across different settings; single time point (48h) versus full time series, and standard cell properties versus high-dimensional features, demonstrates that combining time series data with high-dimensional features is optimal. A single time point may overlook transient states, while a full time series captures critical events essential for MOA prediction. Time series analysis also benefits from noise reduction through filtering, enhancing accuracy. High-dimensional features capture relevant information not found in standard cell properties. These results were consistent in both cell lines.

This underscores the notion that from label-free live-cell imaging, we can strategically extract information-rich features to predict the MOA of a compound. Beyond screening, this method is expected to be broadly useful for studying dynamic cellular responses, providing insight into how cells react to biological and chemical perturbations over time. It has potential to become one important component in phenotyping for next-generation drug discovery.

Data and code availability

All data and code will be made publicly available upon peer-reviewed publication.

ACKNOWLEDGMENTS

O.S. acknowledges funding from the Swedish Research Council (2024-04576, and 2024-03566), FORMAS (Grant 2022-00940), Swedish Cancer Foundation (22 2412 Pj 03 H), and Horizon Europe Grant Agreement 101057442 (REMEDI4ALL). Sartorius has supported this work by providing lab equipment and materials.

AUTHOR CONTRIBUTIONS

Conceptualization, E.F., G.L., J.T., and P.J.; methodology, E.F. and P.J.; investigation, E.F., J.R., G.L., D.H., M.M.J. and P.J.; writing-–original draft, E.F., J.R., G.L., D.H., O.S., and P.J.; writing–review & editing, J.C.P. and J.T.; funding acquisition, O.S. and J.T.; resources, G.L, J.C.P., J.R., and D.H.; supervision, P.J., O.S., J.C.P., G.L., and J.T.

DECLARATION OF INTERESTS

P.J., J.T., and G.L. are employed by Sartorius. J.C.P. and O.S. declare ownership in Phenaros Pharmaceuticals.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES

During the preparation of this work, the author(s) used ChatGPT in order to improve readability and grammar. After using this tool or service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

SUPPLEMENTAL INFORMATION INDEX

Supplementary Figure 1 and Supplementary Tables 1-2 and their legends in a PDF

Supplementary Table 3 Compound concentration information in a csv file

MAIN FIGURE TITLES AND LEGENDS

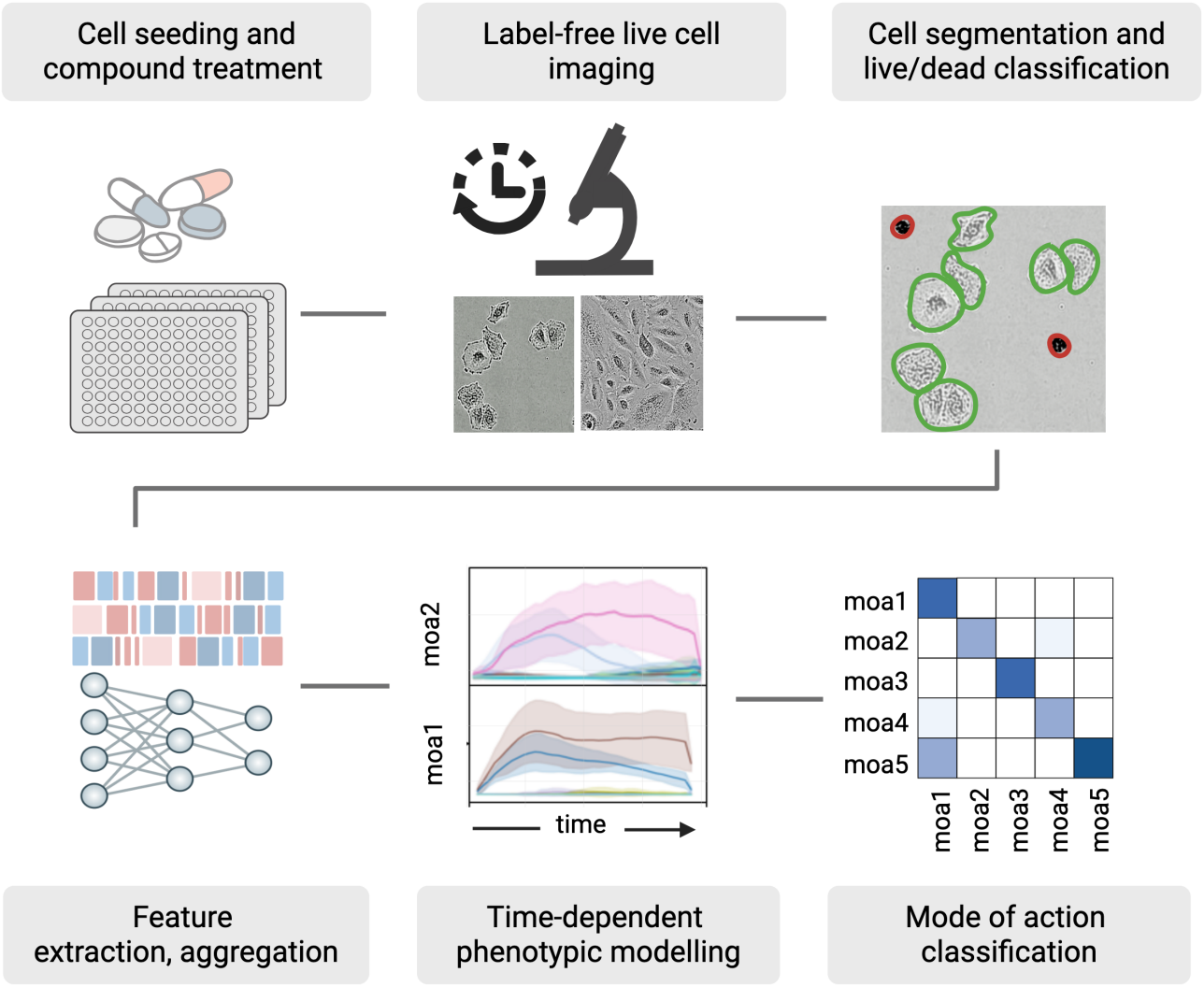

Figure 1: Schematic workflow of the time-dependent phenotypic profiling approach. Cells are seeded and treated with compounds in a multi-well plate format, followed by live-cell imaging at multiple time points for 72 hours. Images are then used to segment single cells and to classify live and dead cells. Single-cell features are extracted using a pre-trained DINOv2-ViT-B14 model, after which the features are aggregated to median well features. These features are used in supervised, time-dependent phenotypic modelling to capture dynamic cellular responses. The resulting Eq. profiles provide insights into temporal effects and are used for MOA classification.

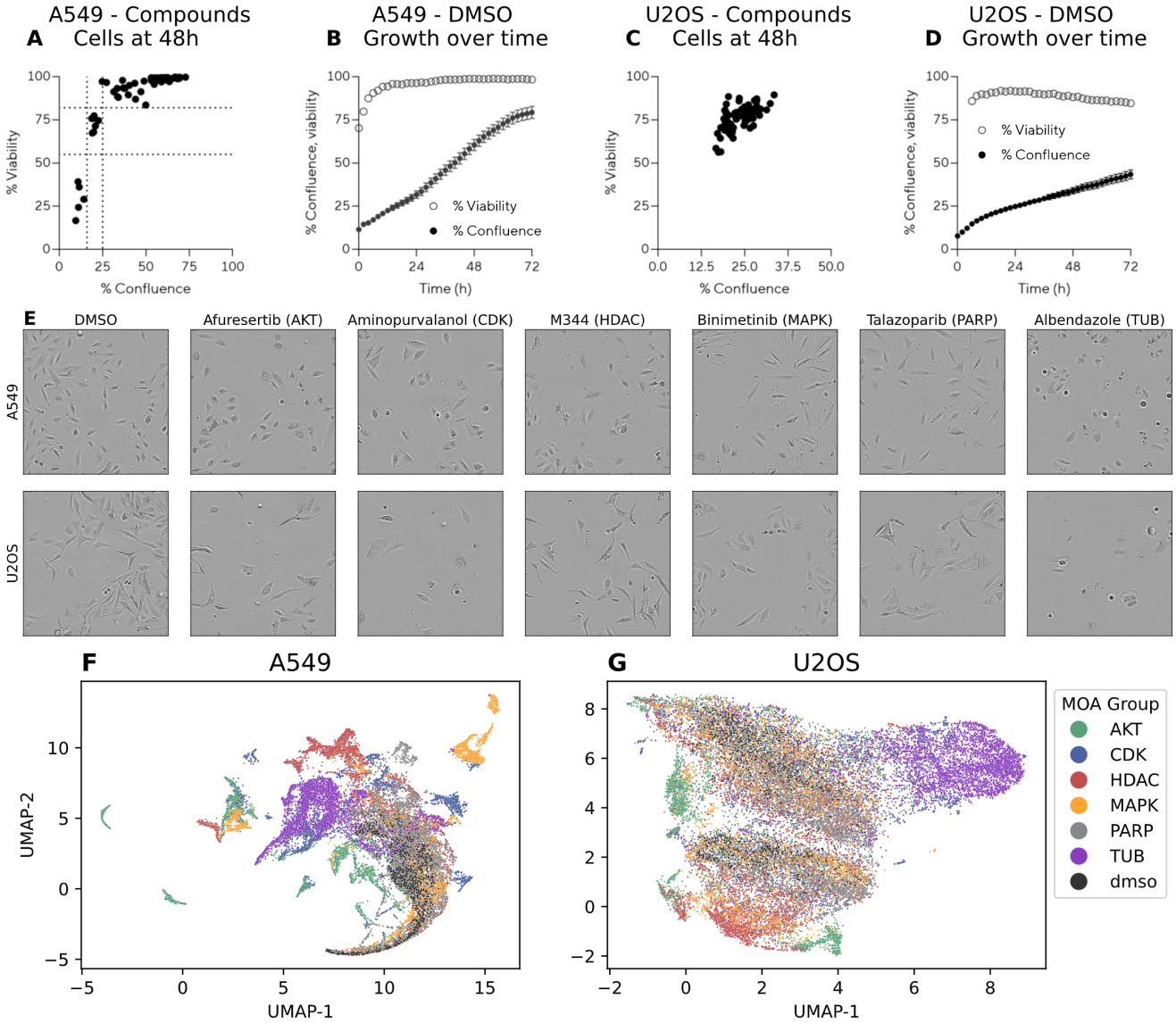

Figure 2: Scatter plots, A549 in A and U2OS in C, indicates the mean % confluence and % viability values of each compound at 48h. Time series plot shows the % confluence and % viability of A549 cells in B and U2OS cells in D in the negative control (DMSO) wells over time. The mean value comprising all negative control wells from all plates is displayed and error bars represent standard deviation. In E, phase contrast images show negative control and compound-treated A549 and U2OS cells exemplifying each mode of action at 48h. In F and G, UMAPs for A549 and U2OS cells including data from every other hour from 24h to 72h.

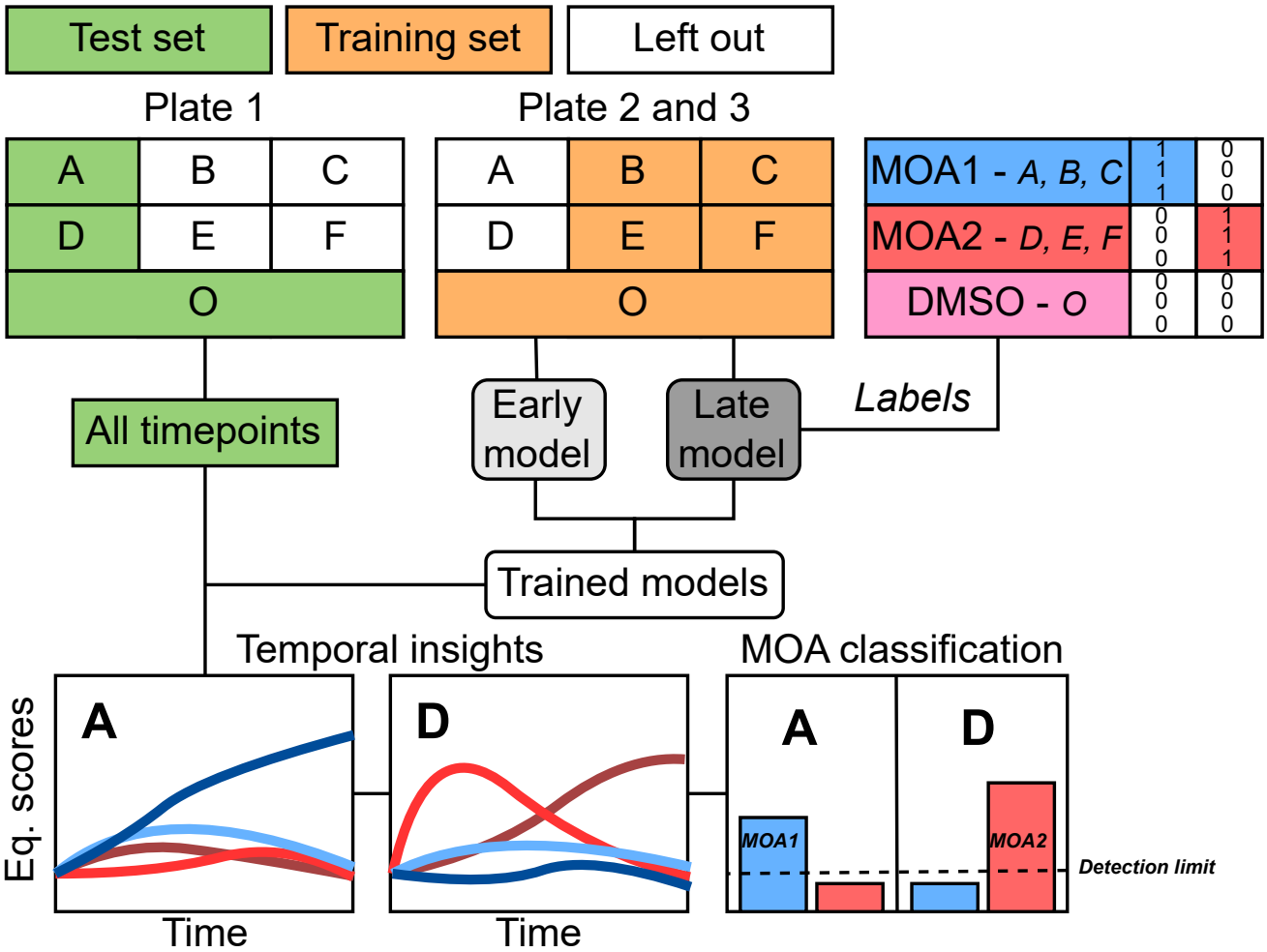

Figure 3: Illustration of a CV round in the data analysis workflow. Compounds A, B, and C represent MOA group 1, while D, E, and F represent MOA group 2. The letter O denotes negative DMSO controls. Green indicates the test set, which is predicted by the model after it has been trained on the orange training set. White represents data excluded from the current CV round. Blue, red, and pink labels MOA1, MOA2, and DMSO, respectively. For the training phase, data (orange) from early (hours 6-14) and late (hours 46-54) time points are used to train two separate models, both utilizing the same labels. The test set is then predicted across all time points (hours 0-72), resulting in four predictions per time point: one early and one late for each MOA. The early and late Eq. scores for the same MOA are smoothed over the time course to mitigate biological variability. These smoothed values are then summed for each MOA per observation. The MOA with maximum sum is used for classification. The negative controls' Eq. scores are used to establish a detection limit for each MOA, which helps filter out false positives and increases confidence in the predictions. This detection limit improves the reliability of the predictions by reducing the number of false positives and enhancing the overall confidence in the results.

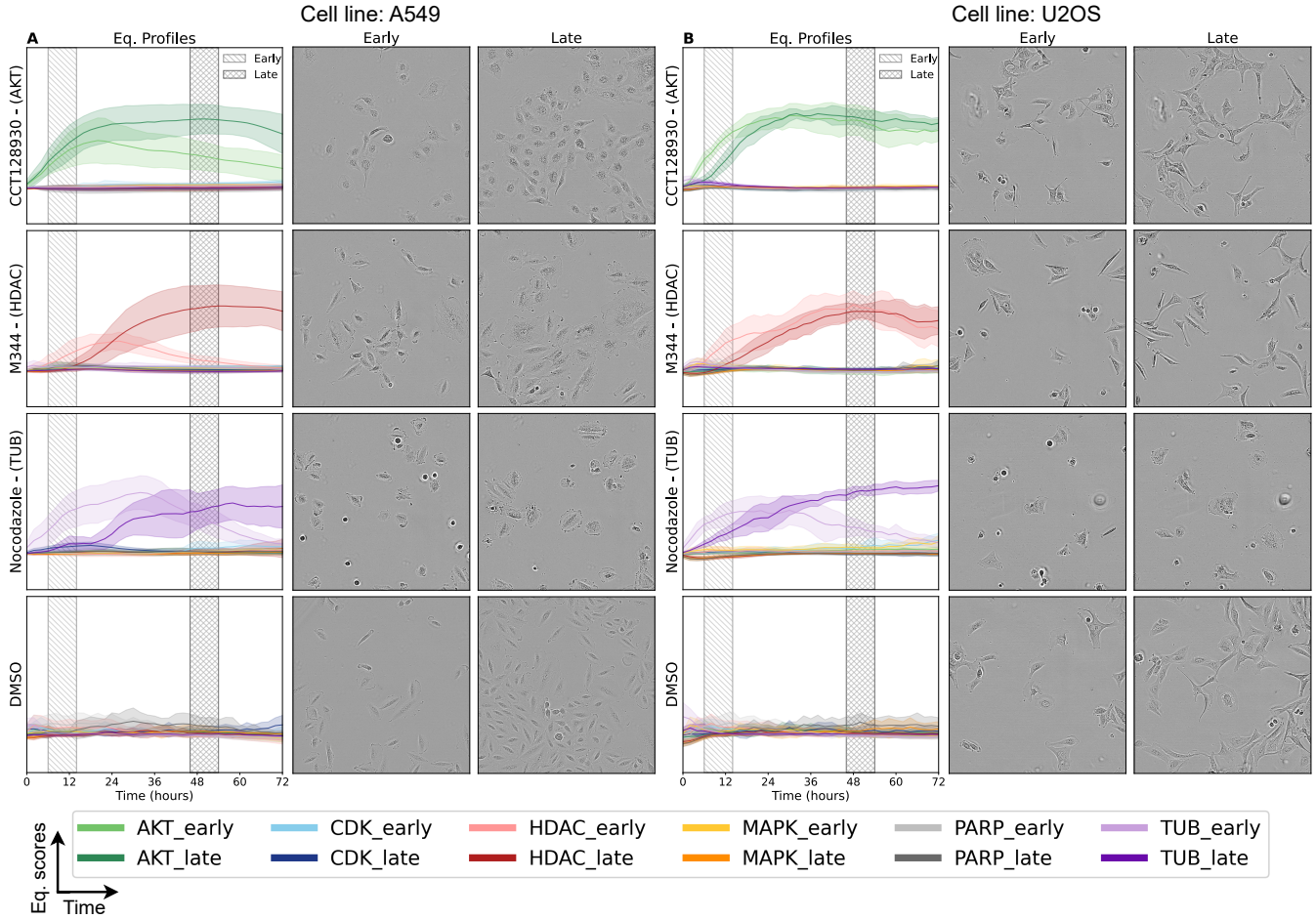

Figure 4: Eq. profiles and live-cell phase contrast images of A549 cells (A) and U2OS cells (B) at 8h and 48h. The first row presents data for CCT128930 (an AKT inhibitor), the second for M344 (an HDAC inhibitor), the third for Nocodazole (a TUB inhibitor), and the fourth for negative controls, DMSO, i.e. untreated cells.

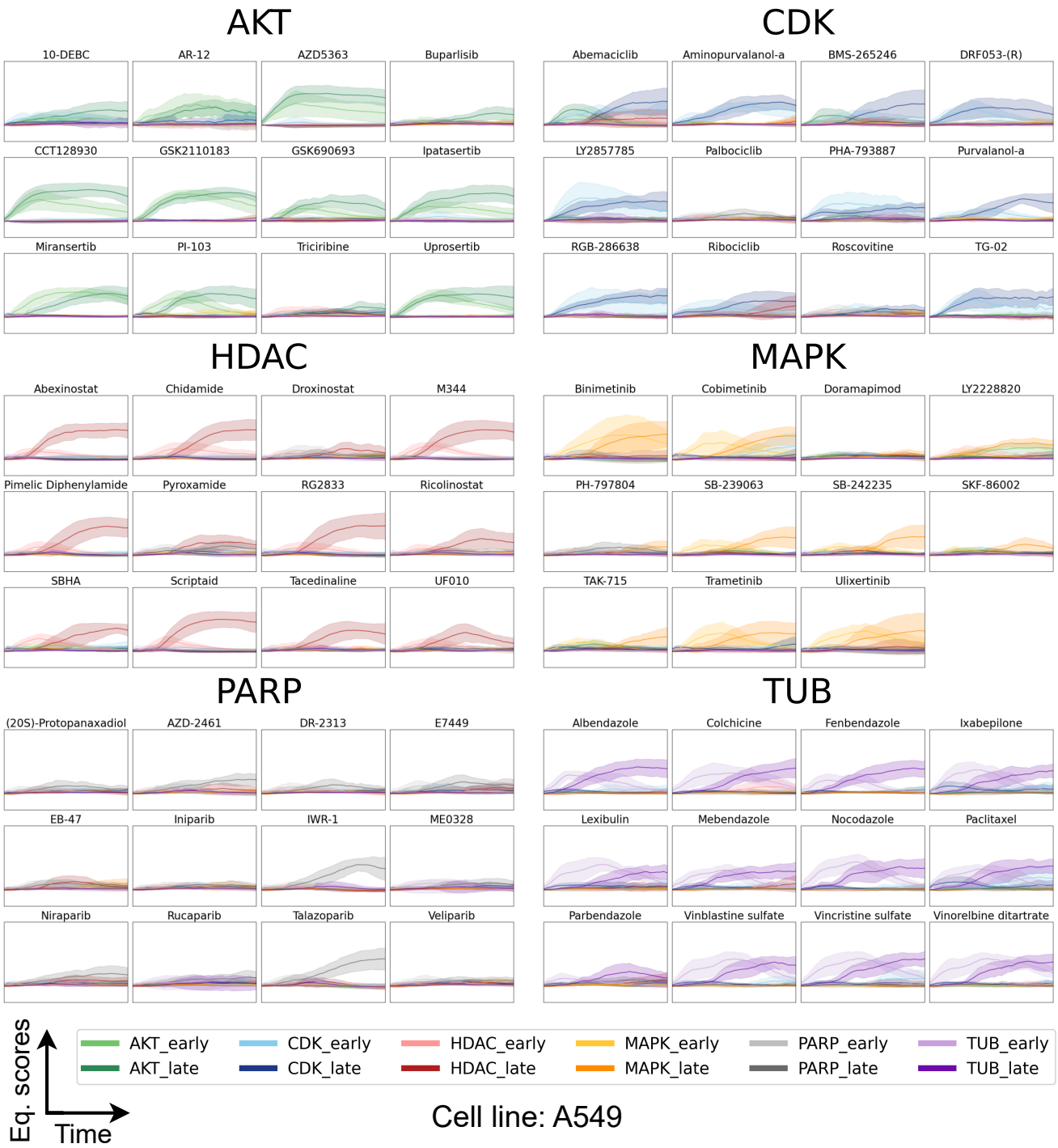

Figure 5: Eq. profiles for the compounds in the dataset grouped by their MOA. All axes have the same scale. The Y-axis represents the Eq. score predicted by the early and late models ranging from -0.5 to 2, while the X-axis shows time, ranging from 0 to 72 hours. Each line in the plot represents the mean Eq. profile from 15 wells across three plates, with one standard deviation depicted as the confidence interval. Consequently, each plot contains twelve Eq. profiles, with two profiles (early and late) for each MOA. Notably, in most cases, only 1-3 profiles are visible, as the others are predicted to be close to zero. Lighter colours indicate early Eq. profiles, whereas darker colours represent late Eq. profiles.

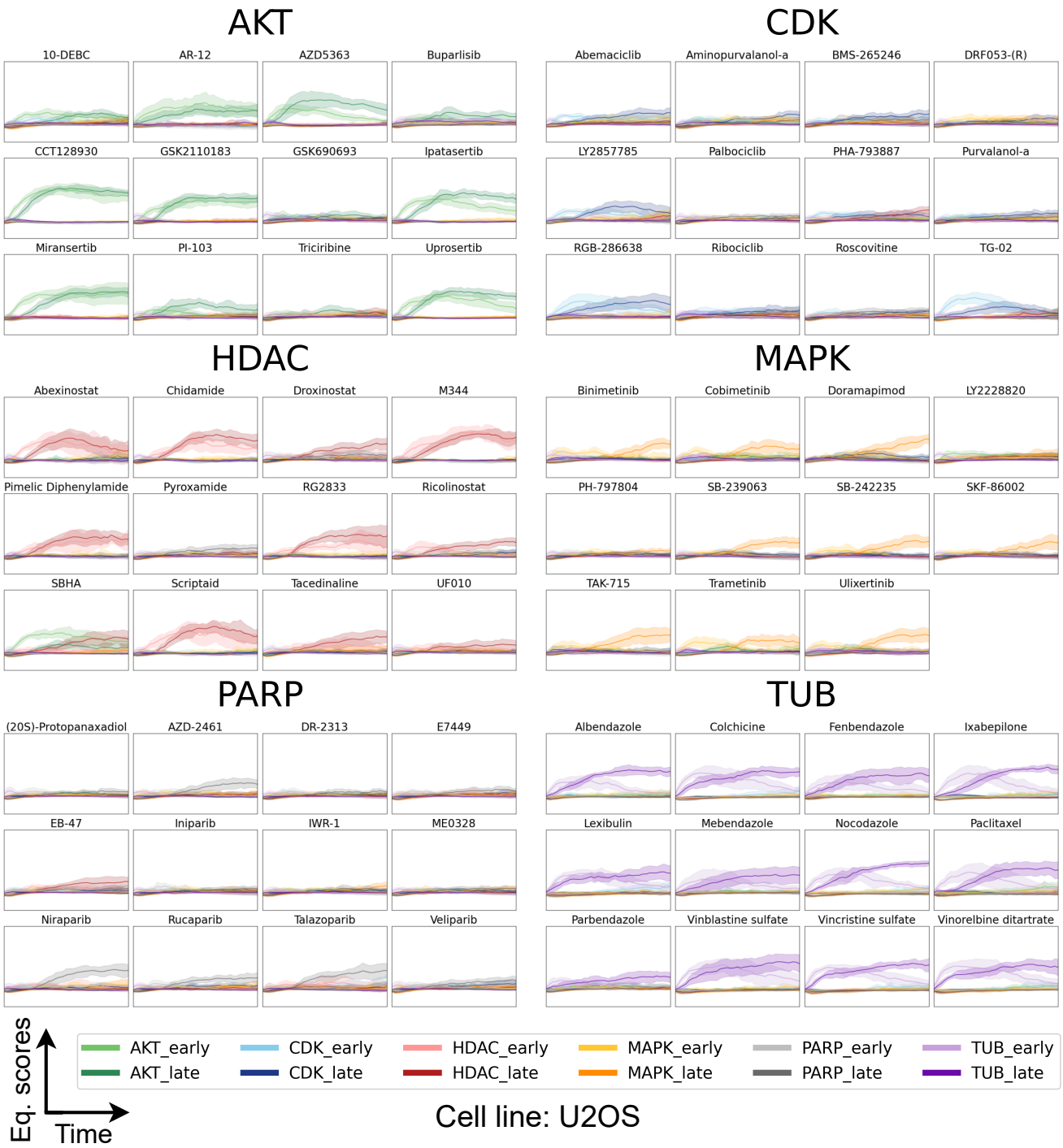

Figure 6: Eq. profiles for the compounds in the dataset grouped by their MOA. All axes have the same scale. The Y-axis represents the Eq. score predicted by the early and late models ranging from -0.5 to 2, while the X-axis shows time, ranging from 0 to 72 hours. Each line in the plot represents the mean Eq. profile from 15 wells across three plates, with one standard deviation depicted as the confidence interval. Consequently, each plot contains twelve Eq. profiles, with two profiles (early and late) for each MOA. Notably, in most cases, only 1-3 profiles are visible, as the others are predicted to be close to zero. Lighter colours indicate early Eq. profiles, whereas darker colours represent late Eq. profiles.

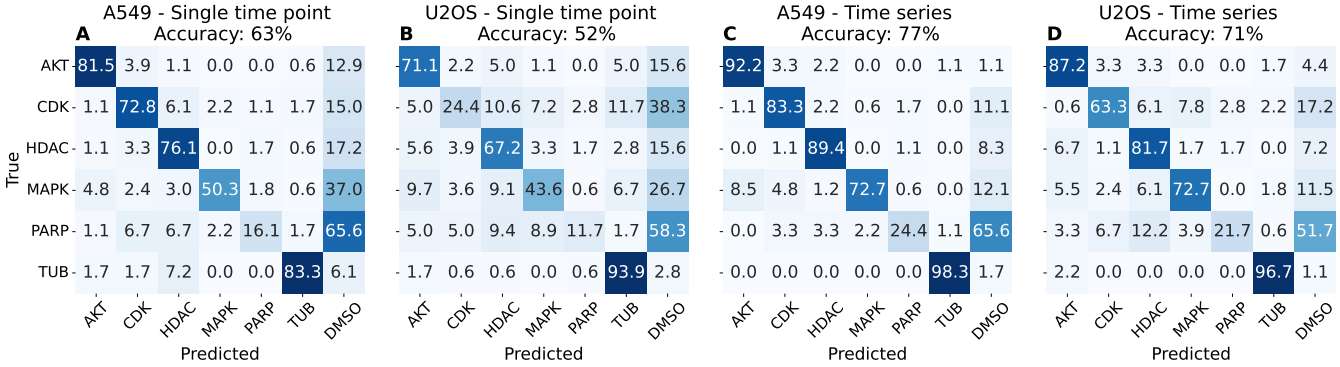

Figure 7: Confusion matrices of classification results using DINO features with the true labels on the Y-axis and the predicted on the X-axis. The data is normalized row-wise. Thus, each number represents the percentage of samples from a given true MOA that are classified into each predicted MOA group. In A and B, the results from single time point (48h) classifications for A549- and U2OS cells with 63% and 52% accuracy respectively are shown. Both show limited performance in classifying MAPK and PAPR correctly and for U2OS, CDK is also hard. In C and D the results from time series classification for A549- and U2OS cells with 77% and 71% accuracy are shown. Here all MOAs are predicted clearly better compared to A and B. Especially MAPK going 50.3% accuracy in A to 72.7% in C for A549 cells and 43.6% in B to 72.7% in D. Observations are classified as negative controls (DMSO) if they are below the detection limit. Negative control observations are not included in the confusion matrix since they would by default be correct ∼97.5% of the time due to the definition of the detection limit.

References

1. Pylvänäinen, J.W., de Mariscal, E.G., Henriques, R., and Jacquemet, G. (2023). Live-cell imaging in the deep learning era. Current Opinion in Cell Biology *85*, 102271. URL: https://www.sciencedirect.com/science/article/pii/S0955067423001205. doi: https://doi.org/10.1016/j.ceb.2023.102271.

2. Bagheri, N., Carpenter, A.E., Lundberg, E., Plant, A.L., and Horwitz, R. (2022). The new era of quantitative cell imaging—challenges and opportunities. Molecular cell *82*, 241–247.

3. Alieva, M., Wezenaar, A.K.L., Wehrens, E.J., and Rios, A.C. (2023). Bridging live-cell imaging and next-generation cancer treatment. Nature Reviews Cancer *23*, 731–745. URL: https://doi.org/10.1038/s41568-023-00610-5. doi: 10.1038/s41568-023-00610-5.

4. Tang, Q., Ratnayake, R., Seabra, G., Jiang, Z., Fang, R., Cui, L., Ding, Y., Kahveci, T., Bian, J., Li, C., Luesch, H., and Li, Y. (2024). Morphological profiling for drug discovery in the era of deep learning. Briefings in Bioinformatics *25*, bbae284. URL: https://doi.org/10.1093/bib/bbae284. doi: 10.1093/bib/bbae284. arXiv:https://academic.oup.com/bib/article-pdf/25/4/bbae284/58244094/bbae284.pdf.

5. Schneidewind, T., Brause, A., Pahl, A., Burhop, A., Mejuch, T., Sievers, S., Waldmann, H., and Ziegler, S. (2020). Morphological profiling identifies a common mode of action for small molecules with different targets. Chembiochem *21*, 3197–3207. doi: 10.1002/cbic.202000381.

6. Bray, M.A., Singh, S., Han, H., Davis, C.T., Borgeson, B., and Carpenter, A.E. (2016). Cell painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. Nature Protocols *11*, 1757–1774. doi: 10.1038/nprot.2016.105.

7. Seal, S., Trapotsi, M.A., Spjuth, O., Singh, S., Carreras-Puigvert, J., Greene, N., Bender, A., and Carpenter, A.E. (2025). Cell painting: a decade of discovery and innovation in cellular imaging. Nature Methods *22*, 254–268. doi: 10.1038/s41592-024-02528-8.

8. Pahl, A., and Sievers, S. (2019). The cell painting assay as a screening tool for the discovery of bioactivities in new chemical matter. In Systems Chemical Biology vol. 1888 of *Methods in Molecular Biology* pp. 115–126.. Humana Press, New York, NY pp. 115–126. doi: 10.1007/978-1-4939-8891-4_6.

9. Rohban, M.H., Fuller, A.M., Tan, C., Goldstein, J.T., Syangtan, D., Gutnick, A., DeVine, A., Nijsure, M.P., Rigby, M., Sacher, J.R. et al. (2022). Virtual screening for small-molecule pathway regulators by image-profile matching. Cell systems *13*, 724–736.

10. Milo, R., and Phillips, R. (2015). Cell Biology by the Numbers. 1st ed.. Garland Science. URL: https://doi.org/10.1201/9780429258770. doi: 10.1201/9780429258770.

11. Bandey, I.N., Adolacion, J.R.T., Romain, G., Paniagua, M.M., An, X., Saeedi, A., Liadi, I., You, Z., Rajanayake, R.B., Hwu, P., Singh, H., Cooper, L.J., and Varadarajan, N. (2021). Designed improvement to t-cell immunotherapy by multidimensional single cell profiling. Journal for Immunotherapy of Cancer *9*, e001877. doi: 10.1136/jitc-2020-001877.

12. Dekkers, J.F., Alieva, M., Cleven, A., Keramati, F., Wezenaar, A.K.L., van Vliet, E.J., Puschhof, J., Brazda, P., Johanna, I., Meringa, A.D., Rebel, H.G., Buchholz, M.B., Barrera Román, M., Zeeman, A.L., de Blank, S., Fasci, D., Geurts, M.H., Cornel, A.M.,

Driehuis, E., Millen, R., Straetemans, T., Nicolasen, M.J.T., Aarts-Riemens, T., Ariese, H.C.R., Johnson, H.R., van Ineveld, R.L., Karaiskaki, F., Kopper, O., Bar-Ephraim, Y.E., Kretzschmar, K., Eggermont, A.M.M., Nierkens, S., Wehrens, E.J., Stunnenberg, H.G., Clevers, H., Kuball, J., Sebestyen, Z., and Rios, A.C. (2023). Uncovering the mode of action of engineered t cells in patient cancer organoids. Nature Biotechnology *41*, 60–69. doi: 10.1038/s41587-022-01397-w.

13. Freckmann, E.C., Sandilands, E., Cumming, E., Neilson, M., Álvaro Román-Fernández, Nikolatou, K., Nacke, M., Lannagan, T.R.M., Hedley, A., Strachan, D., Salji, M., Morton, J.P., McGarry, L., Leung, H.Y., Sansom, O.J., Miller, C.J., and Bryant, D.M. (2022). Traject3d allows label-free identification of distinct co-occurring phenotypes within 3d culture by live imaging. Nature Communications *13*, 5317. URL: https://doi.org/10.1038/s41467-022-32958-x. doi: 10.1038/s41467-022-32958-x.

14. Shroff, H., Testa, I., Jug, F., and Manley, S. (2024). Live-cell imaging powered by computation. Nature Reviews Molecular Cell Biology *25*, 443–463. URL: https://doi.org/10.1038/s41580-024-00702-6. doi: 10.1038/s41580-024-00702-6. Epub 2024 Feb 20.

15. Harrison, P.J., Gupta, A., Rietdijk, J., Wieslander, H., Carreras-Puigvert, J., Georgiev, P., Wählby, C., Spjuth, O., and Sintorn, I.M. (2023). Evaluating the utility of brightfield image data for mechanism of action prediction. PLOS Computational Biology *19*, 1–19. URL: https://doi.org/10.1371/journal.pcbi.1011323. doi: 10.1371/journal.pcbi.1011323.

16. Edlund, C., Jackson, T.R., Khalid, N., Bevan, N., Dale, T., Dengel, A., Ahmed, S., Trygg, J., and Sjögren, R. (2021). Livecell—a large-scale dataset for label-free live cell segmentation. Nature Methods *18*, 1038–1045. doi: 10.1038/s41592-021-01249-6.

17. Ker, D.F.E., Eom, S., Sanami, S., Bise, R., Pascale, C., Yin, Z., Huh, S.i., Osuna-Highley, E., Junkers, S.N., Helfrich, C.J., Liang, P.Y., Pan, J., Jeong, S., Kang, S.S., Liu, J., Nicholson, R., Sandbothe, M.F., Van, P.T., Liu, A., Chen, M., Kanade, T., Weiss, L.E., and Campbell, P.G. (2018). Phase contrast time-lapse microscopy datasets with automated and manual cell tracking annotations. Scientific Data *5*, 180237. URL: https://doi.org/10.1038/sdata.2018.237. doi: 10.1038/sdata.2018.237.

18. Baar, S., Kuragano, M., Tokuraku, K., and Watanabe, S. (2022). Towards a comprehensive approach for characterizing cell activity in bright-field microscopic images. Scientific Reports *12*, 16884. URL: https://doi.org/10.1038/s41598-022-20598-6. doi: 10.1038/s41598-022-20598-6.

19. Wiggins, L., Lord, A., Murphy, K.L., Lacy, S.E., O'Toole, P.J., Brackenbury, W.J., and Wilson, J. (2023). The cellphe toolkit for cell phenotyping using time-lapse imaging and pattern recognition. Nature Communications *14*, 1854. URL: https://doi.org/10.1038/s41467-023-37447-3. doi: 10.1038/s41467-023-37447-3.

20. Doron, M., Moutakanni, T., Chen, Z.S., Moshkov, N., Caron, M., Touvron, H., Bojanowski, P., Pernice, W.M., and Caicedo, J.C. (2023). Unbiased single-cell morphology with self-supervised vision transformers. bioRxiv. URL: https://www.biorxiv.org/content/early/2023/06/18/2023.06.16.545359. doi: 10.1101/2023.06.16.545359. arXiv:https://www.biorxiv.org/content/early/2023/06/18/2023.06.16.545359.full.pdf.

21. Kim, V., Adaloglou, N., Osterland, M., Morelli, F.M., Halawa, M., König, T., Gnutt, D., and Zapata, P.A.M. (2024). Self-supervision advances morphological profiling by unlocking powerful image representations. bioRxiv. URL: https://www.biorxiv.

org/content/early/2024/01/03/2023.04.28.538691. doi: 10.1101/2023.04.28.538691. arXiv:https://www.biorxiv.org/content/early/2024/01/03/2023.04.28.538691.full.pdf.

22. Sartorius (n.d.). Incucyte® ai cell health analysis software module. https://www.sartorius.com/en/products/live-cell-imaging-analysis/live-cell-analysis-software/incucyte-ai-cell-health-analysis-software-module. . Accessed: 10 March 2025.

23. Lovell, G., Porto, D.A., Trigg, J., Holtz, N., Bevan, N., Dale, T., and Appledorn, D. (2023). Ai-driven image analysis enables simplified, label-free cytotoxicity screening. In Proceedings of the American Association for Cancer Research Annual Meeting 2023; Part 1 (Regular and Invited Abstracts) vol. 83. Orlando, FL: AACR pp. Abstract nr 5418.

24. Francisco Rodríguez, M.A., Carreras Puigvert, J., and Spjuth, O. (2023). Designing microplate layouts using artificial intelligence. Artificial Intelligence in the Life Sciences *3*, 100073. URL: https://www.sciencedirect.com/science/article/pii/S266731852300017X. doi: https://doi.org/10.1016/j.ailsci.2023.100073.

25. McInnes, L., Healy, J., Saul, N., and Großberger, L. (2018). Umap: Uniform manifold approximation and projection. Journal of Open Source Software *3*, 861. URL: https://doi.org/10.21105/joss.00861. doi: 10.21105/joss.00861.

26. Ansel, J., Yang, E., He, H., Gimelshein, N., Jain, A., Voznesensky, M., Bao, B., Bell, P., Berard, D., Burovski, E., Chauhan, G., Chourdia, A., Constable, W., Desmaison, A., DeVito, Z., Ellison, E., Feng, W., Gong, J., Gschwind, M., Hirsh, B., Huang, S., Kalambarkar, K., Kirsch, L., Lazos, M., Lezcano, M., Liang, Y., Liang, J., Lu, Y., Luk, C.K., Maher, B., Pan, Y., Puhrsch, C., Reso, M., Saroufim, M., Siraichi, M.Y., Suk, H., Zhang, S., Suo, M., Tillet, P., Zhao, X., Wang, E., Zhou, K., Zou, R., Wang, X., Mathews, A., Wen, W., Chanan, G., Wu, P., and Chintala, S. (2024). Pytorch 2: Faster machine learning through dynamic python bytecode transformation and graph compilation. In Proceedings of the 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2. ASPLOS '24. New York, NY, USA: Association for Computing Machinery. ISBN 9798400703850 pp. 929–947. URL: https://doi.org/10.1145/3620665.3640366. doi: 10.1145/3620665.3640366.

27. Research, F. (2023). Dinov2. https://github.com/facebookresearch/dinov2. .

28. Darcet, T., Oquab, M., Mairal, J., and Bojanowski, P. (2023). Vision transformers need registers. .

29. Oquab, M., Darcet, T., Moutakanni, T., Vo, H.V., Szafraniec, M., Khalidov, V., Fernandez, P., Haziza, D., Massa, F., El-Nouby, A., Howes, R., Huang, P.Y., Xu, H., Sharma, V., Li, S.W., Galuba, W., Rabbat, M., Assran, M., Ballas, N., Synnaeve, G., Misra, I., Jegou, H., Mairal, J., Labatut, P., Joulin, A., and Bojanowski, P. (2023). Dinov2: Learning robust visual features without supervision. .

30. Forsgren, E., Cloarec, O., Jonsson, P., Lovell, G., and Trygg, J. (2024). A scalable, data analytics workflow for image-based morphological profiles. Chemometrics and Intelligent Laboratory Systems *254*, 105232. URL: https://www.sciencedirect.com/science/article/pii/S0169743924001722. doi: https://doi.org/10.1016/j.chemolab.2024.105232.

31. Roberts, S.W. (1959). Control chart tests based on geometric moving averages. Technometrics *1*, 239–250. URL: http://www.jstor.org/stable/1266443.

32. Gross, S.M., and Rotwein, P. (2015). Akt signaling dynamics in individual cells. Journal of Cell Science *128*, 2509–2519. URL: https://doi.org/10.1242/jcs.168773. doi: 10.1242/jcs.168773. arXiv:https://journals.biologists.com/jcs/article-pdf/128/14/2509/1938819/jcs168773.pd

33. Li, G., Tian, Y., and Zhu, W.G. (2020). The roles of histone deacetylases and their inhibitors in cancer therapy. Frontiers in Cell and Developmental Biology *8*. URL: https://www.frontiersin.org/journals/cell-and-developmental-biology/articles/10.3389/fcell.2020.576946. doi: 10.3389/fcell.2020.576946.

34. Lord, C.J., and Ashworth, A. (2017). Parp inhibitors: Synthetic lethality in the clinic. Science *355*, 1152–1158. URL: https://www.science.org/doi/abs/10.1126/science.aam7344. doi: 10.1126/science.aam7344. arXiv:https://www.science.org/doi/pdf/10.1126/science.aam7344.

35. Lafanechère, L. (2022). The microtubule cytoskeleton: An old validated target for novel therapeutic drugs. Frontiers in Pharmacology *13*. URL: https://www.frontiersin.org/journals/pharmacology/articles/10.3389/fphar.2022.969183. doi: 10.3389/fphar.2022.969183.

36. Advani, D., and Kumar, P. (2024). Uncovering cell cycle dysregulations and associated mechanisms in cancer and neurodegenerative disorders: A glimpse of hope for repurposed drugs. Molecular Neurobiology *61*, 8600–8630. URL: https://doi.org/10.1007/s12035-024-04130-7. doi: 10.1007/s12035-024-04130-7.

37. Knudsen, E.S., and Witkiewicz, A.K. (2017). The strange case of cdk4/6 inhibitors: Mechanisms, resistance, and combination strategies. Trends in Cancer *3*, 39–55. URL: https://doi.org/10.1016/j.trecan.2016.11.006. doi: 10.1016/j.trecan.2016.11.006. Doi: 10.1016/j.trecan.2016.11.006.