Interpretable phenotypic profiling of 3D cellular morphodynamics

Matt De Vries[1,2,3][0000−0002−4098−1611], Reed Naidoo[1][0009−0004−7519−197X], Olga Fourkioti[1][0000−0002−1315−8814], Lucas G. Dent[1][0000−0001−8573−4617], Nathan Curry[2][0000−0001−7642−8036], Christopher Dunsby[2][0000−0001−8782−0885], and Chris Bakal[1,3][0000−0002−0413−6744]

[1] The Institute of Cancer Research, London, UK {matt.devries@icr.ac.uk}
[2] Imperial College London, London, UK
[3] Sentinal4D, London, UK

Abstract. The dynamic 3D shape of a cell acts as a signal of its physiological state, reflecting the interplay of environmental stimuli and intra- and extra-cellular processes. However, there is little quantitative understanding of cell shape determination in 3D, largely due to the lack of data-driven methods that analyse 3D cell shape dynamics. To address this, we have developed `MorphoSense`, an interpretable, variable-length multivariate time series classification (TSC) pipeline based on multiple instance learning (MIL). We use this pipeline to classify 3D cell shape dynamics of perturbed cancer cells and learn hallmark 3D shape changes associated with clinically relevant and shape-modulating small molecule treatments. To show the generalisability across datasets, we apply our pipeline to classify migrating T-cells in collagen matrices and assess interpretability on a synthetic dataset. Across datasets, our pipeline offers increased predictive performance and higher-quality interpretations. To our knowledge, our work is the first to utilise MIL for multivariate, variable-length TSC, focusing on interpretable 3D morphodynamic profiling of biological cells.

Keywords: 4D Cell Shape · Graph Neural Networks · Multiple Instance Learning · Time Series Classification

1 Introduction

Cell morphodynamics is a cell's time-dependent morphology or "shape-shifting ability" [9]. This ability is critical for cell division, migration and differentiation. Dysregulation of cell shape determination underpins the ability of cancer cells to escape primary tumours, enter the bloodstream, and localise in a distant and deadly metastatic site. Indeed, various therapeutic strategies have been proposed to inhibit shape-changing processes in cancer [3].

Historically, cell shape analysis has focused on describing cells' fixed, two-dimensional (2D) shapes by read-outs, such as size, eccentricity, and the number of protrusions. This analysis of 2D cell shape has provided valuable information about signalling states [2], cancer metastasis [29], and drug profiling [6].

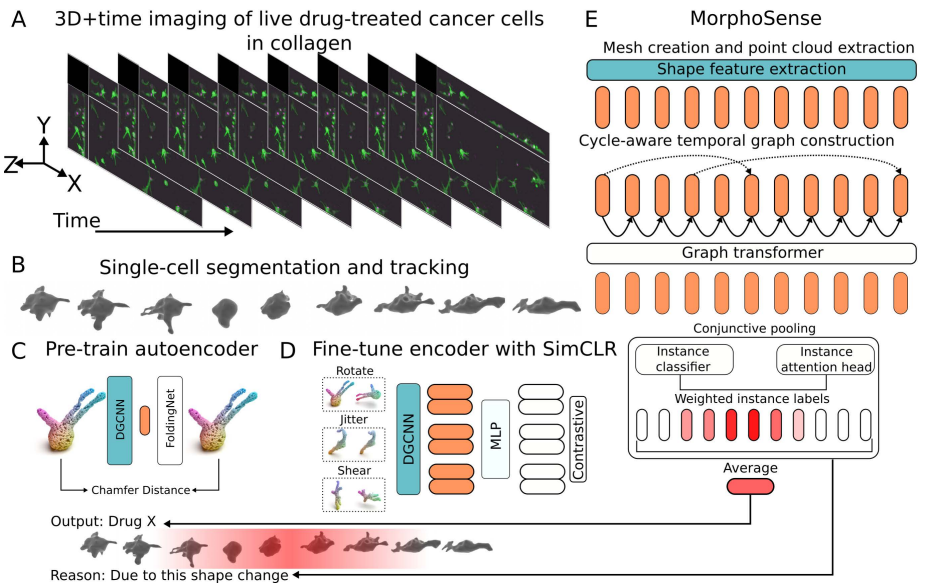

Fig. 1. (**A**) Cells embedded in collagen matrices were imaged in 3D using oblique-plane microscopy. (**B**) Cells and nuclei were segmented and tracked. (**C**) Segmentations were converted to point clouds, and 3D shape features were extracted from each time point. Feature extraction was done by pre-training a DGCNN encoder and (**D**) fine-tuning it with SimCLR. (**E**) 3D shape features were passed through a MIL-based interpretable graph transformer (`MorphoSense`) to predict the label (drug treatment) and output interpretability scores for each time point.

Additionally, live imaging of 2D cell shapes has enabled phenotypic profiling of therapeutics [15, 17]. The recent advancement of three-dimensional (3D) microscopy technologies [22] has facilitated the high-throughput acquisition of 3D cell shapes, paving the way for the deployment of deep learning (DL) techniques to quantify 3D cell shapes effectively. Although examining 3D cell shapes at a single point has the potential to provide important information about the cell state [11], it does not shed light on the dynamic changes cells undergo in response to external stimuli. Understanding those dynamic changes can unveil useful insights about cell shape determination [23, 8]. Specifically, interpretable classification of the dynamic response to pharmacological interventions can offer much information about the underlying drug mechanisms of action and unlock new therapies to combat cancer metastasis. This task can be approached as a multivariate time series classification (TSC) problem. Given a dataset of tracked 3D cells, predicting the drug treatment for each cell based on its cell shape evolution after perturbation in an interpretable manner will allow us to identify the key 3D shape changes associated with different treatments.

Several multivariate TSC algorithms have been proposed in recent years [21]. While these methods have produced state-of-the-art results on popular bench-

mark datasets [1], they often require fixed-length time series - a limitation given that many real-world time series are not uniform in length. To overcome this limitation, long short-term memory models (LSTM), graph-based models [19], and an adaptation to the original Random Convolutional Kernel Transform (ROCKET) [4] have been proposed. However, these methods often lack inherent interpretability. To tackle this issue and produce sparse explanations on time series data, Early et al. (2024) [12] introduced Multiple Instance Learning for Locally Explainable Time (MILLET). Notably, MILLET was the first framework to use multiple instance learning (MIL) on general time series data; however, while a general pipeline, MILLET focused on fixed-length univariate TSC. Similar to ideas presented in [12], we propose applying MIL to TSC due to its inherent interpretability, ease of adaptation, and recent progress in weakly annotated classification tasks [14]. We extend this work by applying MIL to multi-variate variable-length TSC of 3D cell morphodynamics by developing a temporal dependency-aware graph construction and utilising a combination of graph transformers and conjunctive pooling. This has led to the creation of `MorphoSense`, a novel model for interpretable phenotypic profiling of 3D cell morphodynamics (Fig. 1).

The unique contributions of this paper are: (1) an interpretable framework that can be used for multivariate, variable length TSC, (2) a cycle-aware temporal graph construction, (3) the application of this framework to the 3D cellular morphodynamic classification of drug-treated cancer cells in collagen matrices, learning the hallmark shape changes induced by a clinically relevant cancer therapy (Palboclib) and small molecules targeting disease-related cell processes, and (4) the application to another biological dataset of T-cells and a synthetic dataset of 3D shape dynamics.

2 Methods

2.1 Imaging and segmentation

Following treatment with four small-molecule inhibitors, we prepared and imaged live WM266.4 melanoma cancer cells in 3D collagen matrices. Cells were imaged using 3D light-sheet oblique plane microscopy (OPM) (Fig. 1 A). We segmented every cell and nuclei using Otsu's thresholding and active contours, respectively. We then tracked the dynamics of each cell using a simple particle-tracking algorithm (Fig. 1 B).

2.2 Feature extraction

DL techniques have transformed 3D shape analysis of everyday objects [30]. Among the most successful applications have been those that utilised point cloud representations of 3D shapes [31]. Recently, these techniques have been shown to generalise well to 3D shape analysis of cancer cells [10]. Therefore, to analyse 3D cell shape dynamics, we obtained smoothed mesh objects from

cell segmented masks using marching cubes and Laplacian smoothing. We then uniformly sample 2048 points from each mesh object. To extract meaningful 3D shape feature representations from individual time points of cells, we first pre-trained a FoldingNet [31] autoencoder with a dynamic graph convolutional network (DGCNN) [28] encoder (Fig. 1 C). This was trained on a dataset of over 90,000 point cloud representations of fixed 3D melanoma cells from previous works [10] and additional datasets. The DGCNN feature extractor with an additional 2-layered MLP module was fine-tuned using SimCLR [7] (Fig. 1 D). In the SimCLR configuration, two different augmentations (of rotation, jitter, shear, flip, zoom) of the same point cloud were fed as input to the SimCLR model. Only the final EdgeConv block and the projection head were fine-tuned during training by minimising the temperature-scaled cross-entropy (defined as the NT-Xent contrastive loss). The trained model was then fixed and used as a feature extractor to produce 3D shape features for each cell and nuclei pair for each time point.

2.3 The `MorphoSense` pipeline

TSC as a MIL problem: MIL aims to classify bags of instances into classes where only bag-level labels are provided. In its simplest binary classification form, a bag is positive if and only if at least one of its instances is positive; otherwise, it is considered negative. This paper framed multivariate, variable-length time series classification as a MIL problem. Each cell's shape is represented as a feature vector evolving over time. The shape dynamics of each cell was represented as a bag of instances $\mathbf{X}_i \in \mathbb{R}^{N_i \times D}$, such that $\mathbf{X}_i = \{\mathbf{x}_i^1, \mathbf{x}_i^2, \ldots, \mathbf{x}_i^{N_i}\}$, with length $N_i > 1$, where each time point $\mathbf{x}_i^n$ is a D-dimensional vector, and D is the number of variables in the time series. A common MIL assumption is that instances within a bag are independent and identically distributed. However, this is not the case for time series classification. Therefore, we assume that there are temporal relationships between instances.

Graph construction: Time series data often encapsulate complex dynamics characterised by time-dependent interactions (e.g., delayed effects). Hence, we aimed to construct a graph to encode the direct temporal dependencies between different cellular states over time and the intricate interconnections among multiple shape state cycles across time. In this work, we represented each time series as a weighted attributed graph, $\mathcal{G}_i = (\mathbf{A_i}, \mathbf{X_i})$, consisting of a weighted adjacency matrix $\mathbf{A_i}$ and $\mathbf{A_i} \in \mathbb{R}^{N_i \times N_i}$ and a node-feature matrix $\mathbf{N_i} \in \mathbb{R}^{N_i \times D}$. The adjacency matrix represents the graph topology and can be characterised by the set of V nodes $\mathcal{V} = \{v_1, v_2, \ldots, v_T\}$, and $\mathcal{E} = \{e_{i,j} := (v_i, v_j) \in \mathcal{V} \times \mathcal{V} \mid \mathbf{A}_{i,j} \neq 0\}$, the set of edges, such that $\mathbf{A}_{i,j}$ is the (i,j)-th element in $\mathbf{A}$. The node feature matrix $\mathbf{X}$ contains attributes for each node, where the i-th row $\mathbf{x}_i \in \mathbb{R}^D$ represents the D-dimensional feature vector of node v_i.

When building our graph, we distinguished between two types of edges: temporal and similarity-based. The temporal edges $\mathcal{E}$ are inherently directional and designed to capture the temporal progression from one cellular state to the next. They connect each time point t to the next one $t+1$, such that each edge

$e_{t,t+1} \in \mathcal{E}$ represents a temporal connection from time point t to time point $t+1$, for all $t \in \{1, 2, \ldots, N_i - 1\}$. Thus, the set of temporal edges can be defined as $\mathcal{E} = \{(t, t+1)\, i < N_i\}$.

The similarity-based edges aimed to encapsulate the relationship between different time segments based on their feature representations. For a given time series $\mathbf{X}_i = \{\mathbf{x}_i^1, \mathbf{x}_i^2, \ldots, \mathbf{x}_i^{N_i}\}$, representing different cell shapes over distinct points in time, we constructed a similarity matrix S where each element s_{tm} denotes the cosine similarity between nodes t and m (where $t \neq m$ thereby excluding self-loops). For the formation of weighted edges, a threshold θ is applied to ensure that an edge e_{tm} is added to the graph if $s_{tm} > \theta$. The weight of each edge e_{tm} was set to the similarity value s_{tm}.

The final weighted adjacency matrix, $\mathbf{A}$ of each time series $\mathbf{X}_i$, we had $\mathbf{A}_{t,t+1} = 1$ for all $t < N_i$, and $\mathbf{A}_{t,m} = s_{tm} > \theta$ for all $t, m \in N_i$.

Interpretable Graph Transformer: Given a graph representation of a time series, we aimed to classify individual nodes and the graph. To this end, we combined graph transformer layers [26] with conjunctive pooling [12], dubbed `MorphoSense` (Fig 1. E). We first passed the weighted graph through two graph transformer layers to capture the temporal and cycle relationships among features. Specifically, given the node features (multivariate time series)[4], $\mathbf{X} = \{\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^N\}$, we calculated multi-head attention for each edge from m to n following [26] directly:

$$\begin{aligned} \mathbf{Q}_{c,n} &= \mathbf{W}_{c,q}\mathbf{x}^n + \mathbf{b}_{c,q} \\ \mathbf{K}_{c,m} &= \mathbf{W}_{c,k}\mathbf{x}^m + \mathbf{b}_{c,k} \\ e_{c,nm} &= \mathbf{W}_{c,e}e_{nm} + \mathbf{b}_{c,e} \\ \alpha_{c,nm} &= \frac{\langle \mathbf{Q}_{c,n}, \mathbf{K}_{c,m} + e_{c,nm} \rangle}{\sum_{u \in \mathcal{N}(n)} \langle \mathbf{Q}_{c,n}, \mathbf{K}_{c,u} + e_{c,nu} \rangle} \end{aligned} \tag{1}$$

where $\langle q, k \rangle = \exp\left(\frac{q^T k}{\sqrt{d}}\right)$ is the exponential dot product and d is the hidden size of each head, c is the the attention head. For the c-th attention head, source feature vector $\mathbf{x}^n$ and distant feature vector $\mathbf{x}^m$ are transformed in query matrix $\mathbf{Q}_{c,n} \in \mathbb{R}^{D \times d}$ and key matrix $\mathbf{K}_{c,m} \in \mathbb{R}^{D \times d}$ using trainable weights $\mathbf{W}_{c,q}$, $\mathbf{W}_{c,k}$ and biases $\mathbf{b}_{c,q}$, $\mathbf{b}_{c,k}$. The edge features are encoded by $\mathbf{W}_{c,e}$ and $\mathbf{b}_{c,e}$, which is added to the key matrix. After obtaining the graph multi-head attention, a message aggregation is performed from distant m to n:

$$\begin{aligned} \mathbf{V}_{c,m} &= \mathbf{W}_{c,v}\mathbf{x}^m + \mathbf{b}_{c,v} \\ \hat{\mathbf{x}}^{n_{out}} &= \|_{c=1}^{C} \left[\sum_{m \in \mathcal{N}(n)} \alpha_{c,nm} \left(\mathbf{V}_{c,m} + e_{c,nm} \right) \right] \end{aligned} \tag{2}$$

where $\|$ is the concatenation operation for C head attention. Following Shi et al. (2021) [26], we also used a gated residual connection between layers to give

[4] We drop the subscript for ease of notation here.

an output of the graph transformer:

$$\begin{aligned} \mathbf{r}_n &= \mathbf{W}_r \mathbf{x}^n + \mathbf{b}_r \\ \beta_n &= \text{sigmoid}\left(\mathbf{W}_g \left[\hat{\mathbf{x}}^{n_{out}}; \mathbf{r}_n; \hat{\mathbf{x}}^{n_{out}} - \mathbf{r}_n\right]\right) \\ \mathbf{x}^{n_{out}} &= \text{ReLU}\left(\text{LayerNorm}\left(\left(1-\beta_n\right)\hat{\mathbf{x}}^{n_{out}} + \beta_n \mathbf{r}_n\right)\right) \end{aligned} \tag{3}$$

The graph transformer's output was directed into a conjunctive pooling module [12]. This dual-stream architecture consists of a classification head and an attention head. The classification head in a linear layer that generates instance logits $\hat{\mathbf{y}}_i$, while the attention head is responsible for generating an attention score for each instance of the node embedding, enabling a weighted summary of its features based on their importance. The attention scores for each time point, n, were computed according to the following:

$$a_n = \text{sigmoid}\left(\mathbf{W}_2 \cdot \tanh\left(\mathbf{W}_1 \mathbf{x}^{n_{out}} + \mathbf{b}_1\right) + \mathbf{b}_2\right), \tag{4}$$

where $\mathbf{W}_1 \in \mathbb{R}^{8\times 256}$ and $\mathbf{W}_2 \in \mathbb{R}^{1\times 8}$ are trainable weight matrices with corresponding bias vectors, $\mathbf{b}_1 \in \mathbb{R}^8$ and $\mathbf{b}_2 \in \mathbb{R}^1$, and $\tanh(\cdot)$ is the hyperbolic tangent function.

By maintaining parallel processing, the model leverages the strengths of instance-level classification accuracy and attention-driven importance weighting. The final output was produced by scaling the instance logits $\hat{\mathbf{y}}$ using the attention scores $\mathbf{a}$:

$$\hat{Y} = \sum_{j=1}^{t} a_j \hat{y}_j. \tag{5}$$

Interpretations for each time point are class-specific importance measures defined as the multiplication of the instance attention and logit.

3 Experiments

We compared `MorphoSense` to other multivariate TSC models that handle variable length time series without explicit padding or cropping (LSTM). We also evaluated against a transformer-based MIL technique typically applied to whole slide image classification (TransMIL [25]) and another graph transformer-based MIL model (GTP [32]). TransMIL utilises variable length positional encoding through PPEG, and GTP relies on graph structure to incorporate temporal information. Other popular MIL models such as DSMIL [20] and ABMIL [18] were left out of our analysis due to the permutation invariance of bags under their framework - a violation of the temporal ordering of time series. We use the same features for all models and the same graph structure for GTP.

We evaluated our model's interpretability using the process described in [13]. This process uses two metrics: Area Over the Perturbation Curve to Random (AOPCR) and Normalised Discounted Cumulative Gain at n (NDCG@n). Both approaches are ranking metrics rather than actual interpretation values. AOPCR

evaluates interpretability without time point labels, and NDCG@n uses time point labels. We evaluated classification accuracy regarding balanced class accuracy (ACC) and area under the receiver operating curve (AUC). For all experiments, we ran 10-fold cross-validation and reported the mean and standard deviation of the metrics.

4 Datasets

We assessed `MorphSense` on three distinct datasets.

Synthetic dynamic shapes: To evaluate the ability of `MorphoSense` to produce high-quality interpretations, we created a synthetic dataset of 3D shapes morphing into other shapes. This dataset consists of 3 classes defined by their shape changes. Each time point is a point cloud representation of a 3D shape, and each time series varies in length (between 15 and 70 time points). We generate spheres with radii between 0.8 and 1.2 and inject one of three shapes (torus, cube, cylinder) at various time points. We then linearly interpolate between a source sphere and the target shape and back to a sphere. This interpolation happens over varying lengths in the dataset. The whole transition is labelled as class-specific instances. The important instances of each time series are known. Therefore, we can directly evaluate the interpretability of each model. The dataset consisted of 1000 point cloud sequences.

Live drug-treated cancer cells: We used `MorphoSense` to learn the small molecule perturbation applied to live melanoma cancer cells. We performed one-vs-the-rest classification for four drugs: Blebbistatin (myosin inhibitor), CK666 (actin polymerisation inhibitor), Palbociclib (CDK4/CDK6 inhibitor), and PF228 (Focal adhesion kinase (FAK) inhibitor) and report the average results across all models. This dataset consisted of 442 sequences of 3D cells ranging from a time series length of 5 to 116.

Migrating T-cells: Finally, we applied our model to classify intravital two-photon microscopy images of T-cells migrating in the popliteal lymph node (LM), submandibular salivary gland (SMG), and skin [23]. Briefly, Medyukhina et al. (2020) [23] segmented and tracked T-cells in 3D and obtained cell surface meshes for each cell for each time point. The dataset consisted of 869 point cloud sequences ranging from 5 to 75 time points per series. We extracted point clouds from these surface meshes and fed these through our pipeline. We used a pre-trained feature extractor trained on fixed WM266.4 3D melanoma cells.

5 Results

Synthetic dynamic shapes: We compared our model's interpretability against that of GTP and TransMIL. GTP uses GraphCAM, which propagates transformer attention maps and class relevance scores through the network and then reconstructs the class activation map using the graph's adjacency matrix. In TransMIL, the attention coefficients correspond to the class token of the transformer attention matrix. `MorphoSense`, which uses the weighted instance logits

Table 1. Inperpretability and classification results. **Left:** Interpretability results (AOPCR and NDCGn) on **Synthetic Dynamic Shapes** dataset. LSTM does not output interpretation. **Middle:** Classification results of different on drug-treated cancer cells (Melanoma). **Right:** Classification results on the T-cells dataset.

	Dynamic Shapes		Melanoma		T-cells	
Method	AOPCR(↑)	NDCG@n(↑)	ACC(↑)	AUC(↑)	ACC(↑)	AUC(↑)
LSTM	-	-	$0.761_{0.078}$	$0.544_{0.004}$	$0.771_{0.031}$	$0.877_{0.032}$
TransMIL	0.639	$0.277_{0.256}$	$0.735_{0.057}$	$0.705_{0.073}$	$0.786_{0.021}$	$\mathbf{0.918}_{0.013}$
GTP	2.236	$0.743_{0.339}$	$0.745_{0.059}$	$0.699_{0.058}$	$0.767_{0.015}$	$0.911_{0.010}$
`MorphoSense`	**5.992**	$\mathbf{0.853}_{0.110}$	$\mathbf{0.764}_{0.056}$	$\mathbf{0.735}_{0.066}$	$\mathbf{0.793}_{0.028}$	$0.912_{0.017}$

from the conjunctive pooling module, outperformed all models in both interpretability metrics.

Biological cells: `MorphoSense` outperformed all models regarding ACC on both datasets and AUC on the melanoma dataset. While TransMIL outperformed `MorphoSense` in AUC on the migrating T-cell dataset by a small margin, it falls short in terms of interpretability. Fig. 2. shows example qualitative interpretations for each drug and how these relate to classical shape features (see supplementary material). Interestingly, the `MorphoSense` model successfully identified salient morphological changes in drug-induced cells. For example, the model assigns high importance to transforming cells from spherical shapes to more protrusive forms, a change induced by Blebbistatin [11, 10]. The model also identified that CK666 [24], which inhibits actin polymerisation [16], results in cells adopting rounder shapes due to their diminished capacity to form protrusions [11]. PF228 inhibits focal adhesion kinase (FAK), and perturbations that increase FAK activity have been shown to promote flatter cell shapes [5]. Palbociclib is a selective cyclin-dependent kinase (CDK)4/6 inhibitor shown to increase size in 2D cell cultures [27].

6 Conclusion

To our knowledge, `MorphoSense` is the first to utilise MIL for interpretable 4D phenotypic profiling. We have applied our model and several others to one synthetic dataset of shape dynamics and two biological datasets of migrating T-cells and cancer cells following treatment by small molecules targeting disease-related cell processes. Across all datasets, our model was superior in classification and interpretability performance. Identifying universal characteristics of shape change is difficult due to the limited number of publicly available 3D+time datasets of cells. Future work will focus on deeper insights into the hallmark 3D shape changes associated with drug treatment and validation on more datasets.

Acknowledgments. This work was funded by a UK Engineering and Physical Sciences Research Council Impact Acceleration grant (EP/K503733/1), a CRUK Mul-

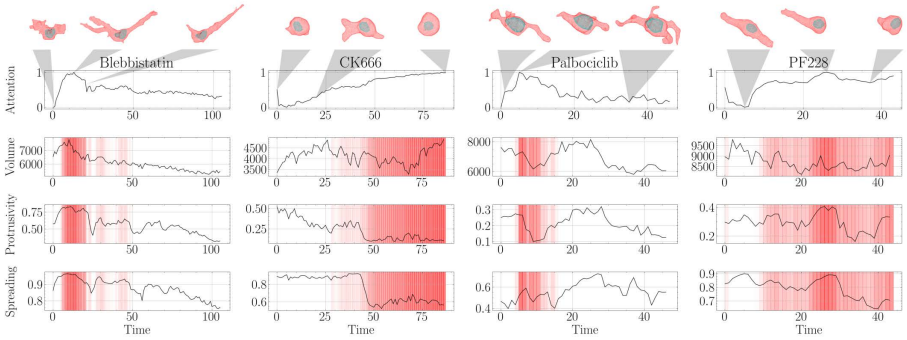

Fig. 2. An example of the interpretations for drug-treated melanoma cells. **Top:** Importance scores for each time point. **2nd row:** Volume of cell across time. **3rd row:** Protrusivity of cell across time. **Bottom:** Spreading of the cell across time.

tidisciplinary Project Award (C53737/A24342), a CRUK-funded Accelerator Award (A29368), and The UK Terry Fox Association.

Disclosure of Interests. C.D. has a licensed granted patent on the optical arrangement for OPM under patent nos. US 8582203 B2 and EP 2316048 B1.

References

1. Bagnall, A., Dau, H.A., Lines, J., Flynn, M., Large, J., Bostrom, A., Southam, P., Keogh, E.: The uea multivariate time series classification archive, 2018 (2018)
2. Bakal, C., Aach, J., Church, G., Perrimon, N.: Quantitative morphological signatures define local signaling networks regulating cell morphology. Science (2007)
3. Barcelo, J., Samain, R., Sanz-Moreno, V.: Preclinical to clinical utility of rock inhibitors in cancer. Trends in Cancer (2023)
4. Bier, A., Jastrzębska, A., Olszewski, P.: Variable-length multivariate time series classification using rocket: A case study of incident detection. IEEE Access (2022)
5. Castillo-Badillo, J.A., Gautam, N.: An optogenetic model reveals cell shape regulation through FAK and fascin. Journal of Cell Science (2021)
6. Chandrasekaran, S.N., Ceulemans, H., Boyd, J.D., Carpenter, A.E.: Image-based profiling for drug discovery: due for a machine-learning upgrade? Nature Reviews Drug Discovery (2021)
7. Chen, T., Kornblith, S., Norouzi, M., Hinton, G.E.: A simple framework for contrastive learning of visual representations. CoRR (2020)
8. Cooper, S., Sadok, A., Bousgouni, V., Bakal, C.: Apolar and polar transitions drive the conversion between amoeboid and mesenchymal shapes in melanoma cells. Molecular Biology of the Cell (2015)
9. Copperman, J., Gross, S.M., Chang, Y.H., Heiser, L.M., Zuckerman, D.M.: Morphodynamical cell state description via live-cell imaging trajectory embedding. Communications Biology (2023)
10. De Vries, M., Dent, L.G., Curry, N., Rowe-Brown, L., Tyson, A., Dunsby, C., Bakal, C.: 3d single-cell shape analysis of cancer cells using geometric deep learning. In: NeurIPS 2022 Workshop on Learning Meaningful Representations of Life (2022)

11. Dent, L.G., Curry, N., Sparks, H., Bousgouni, V., Maioli, V., Kumar, S., Munro, I., Butera, F., Jones, I., Arias-Garcia, M., Rowe-Brown, L., Dunsby, C., Bakal, C.: Environmentally dependent and independent control of 3d cell shape. Cell Reports (2024)
12. Early, J., Cheung, G., Cutajar, K., Xie, H., Kandola, J., Twomey, N.: Inherently interpretable time series classification via multiple instance learning. In: The Twelfth International Conference on Learning Representations (2024)
13. Early, J., Evers, C., Ramchurn, S.: Model agnostic interpretability for multiple instance learning. In: International Conference on Learning Representations (2022)
14. Fourkioti, O., De Vries, M., Bakal, C.: CAMIL: Context-aware multiple instance learning for cancer detection and subtyping in whole slide images. In: The Twelfth International Conference on Learning Representations (2024)
15. Gordonov, S., Hwang, M.K., Wells, A., Gertler, F.B., Lauffenburger, D.A., Bathe, M.: Time series modeling of live-cell shape dynamics for image-based phenotypic profiling. Integrative Biology (2015)
16. Heck, T., Vargas, D.A., Smeets, B., Ramon, H., Van Liedekerke, P., Van Oosterwyck, H.: The role of actin protrusion dynamics in cell migration through a degradable viscoelastic extracellular matrix: Insights from a computational model. PLOS Computational Biology (2020)
17. Heinemann, T., Kornauth, C., Severin, Y., Vladimer, G.I., Pemovska, T., Hadzijusufovic, E., Agis, H., Krauth, M.T., Sperr, W.R., Valent, P., Jäger, U., Simonitsch-Klupp, I., Superti-Furga, G., Staber, P.B., Snijder, B.: Deep Morphology Learning Enhances Ex Vivo Drug Profiling-Based Precision Medicine. Blood Cancer Discovery (2022)
18. Ilse, M., Tomczak, J., Welling, M.: Attention-based deep multiple instance learning. In: Proceedings of the 35th International Conference on Machine Learning. Proceedings of Machine Learning Research (2018)
19. Jin, M., Koh, H.Y., Wen, Q., Zambon, D., Alippi, C., Webb, G.I., King, I., Pan, S.: A survey on graph neural networks for time series: Forecasting, classification, imputation, and anomaly detection. arXiv (2023)
20. Li, B., Li, Y., Eliceiri, K.W.: Dual-stream multiple instance learning network for whole slide image classification with self-supervised contrastive learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2021)
21. Lines, J., Taylor, S., Bagnall, A.: Time series classification with hive-cote: The hierarchical vote collective of transformation-based ensembles. ACM Trans. Knowl. Discov. Data (2018)
22. Maioli, V., Chennell, G., Sparks, H., Lana, T., Kumar, S., Carling, D., Sardini, A., Dunsby, C.: Time-lapse 3-d measurements of a glucose biosensor in multicellular spheroids by light sheet fluorescence microscopy in commercial 96-well plates. Scientific Reports (2016)
23. Medyukhina, A., Blickensdorf, M., Cseresnyés, Z., Ruef, N., Stein, J.V., Figge, M.T.: Dynamic spherical harmonics approach for shape classification of migrating cells. Scientific Reports (2020)
24. Nolen, B.J., Tomasevic, N., Russell, A., Pierce, D.W., Jia, Z., McCormick, C.D., Hartman, J., Sakowicz, R., Pollard, T.D.: Characterization of two classes of small molecule inhibitors of arp2/3 complex. Nature (2009)
25. Shao, Z., Bian, H., Chen, Y., Wang, Y., Zhang, J., Ji, X., Zhang, Y.: TransMIL: Transformer based correlated multiple instance learning for whole slide image classification. In: Beygelzimer, A., Dauphin, Y., Liang, P., Vaughan, J.W. (eds.) Advances in Neural Information Processing Systems (2021)

26. Shi, Y., Huang, Z., Feng, S., Zhong, H., Wang, W., Sun, Y.: Masked label prediction: Unified message passing model for semi-supervised classification. In: Zhou, Z.H. (ed.) Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (2021)
27. Tan, C., Ginzberg, M.B., Webster, R., Iyengar, S., Liu, S., Papadopoli, D., Concannon, J., Wang, Y., Auld, D.S., Jenkins, J.L., Rost, H., Topisirovic, I., Hilfinger, A., Derry, W.B., Patel, N., Kafri, R.: Cell size homeostasis is maintained by cdk4-dependent activation of p38 mapk. Developmental Cell (2021)
28. Wang, Y., Sun, Y., Liu, Z., Sarma, S.E., Bronstein, M.M., Solomon, J.M.: Dynamic graph cnn for learning on point clouds. ACM Transactions on Graphics (TOG) (2019)
29. Wu, P.H., Gilkes, D.M., Phillip, J.M., Narkar, A., Cheng, T.W.T., Marchand, J., Lee, M.H., Li, R., Wirtz, D.: Single-cell morphology encodes metastatic potential. Science Advances (2020)
30. Wu, Z., Song, S., Khosla, A., Yu, F., Zhang, L., Tang, X., Xiao, J.: 3d shapenets: A deep representation for volumetric shapes. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2015)
31. Yang, Y., Feng, C., Shen, Y., Tian, D.: Foldingnet: Point cloud auto-encoder via deep grid deformation. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2018)
32. Zheng, Y., Gindra, R.H., Green, E.J., Burks, E.J., Betke, M., Beane, J.E., Kolachalama, V.B.: A graph-transformer for whole slide image classification. IEEE Transactions on Medical Imaging (2022)