Deep dynamical modelling of drug delivery based on microscopy image data

Philip J Harrison[a,*], Alan Sabirsh[b], Håkan Wieslander[c], Carolina Wählby[c], Johan Karlsson[d], Andreas Hellander[c] & Ola Spjuth[a]

a Department of Pharmaceutical Biosciences
b Advanced Drug Delivery, Pharmaceutical Sciences, R&D, AstraZeneca, Gothenburg
c Department of Information Technology
d Quantitative Biology, Discovery Sciences, R&D, AstraZeneca, Gothenburg

* philip.harrison@farmbio.uu.se

UPPSALA UNIVERSITET

Background

Automated microscopy imaging is one of the most powerful tools to investigate complex biological processes. Rapidly developing high-throughput techniques, capable of generating data at an unprecedented rate, are placing the biomedical sciences on the verge of a digital explosion. Transformative approaches for the analysis of these massive spatio-temporal image datasets are urgently needed.

Dataset

AstraZeneca (AZ) R&D use image based screening systems to study a variety of processes. Current assays tend to focus on single time points and do not interrogate biological systems over time. AZ are currently exploring RNA-based therapeutics. Although such therapies have shown significant promise, more research into RNA delivery is needed before it can transform healthcare. One promising methods of delivery is through Lipid Nano-Particles (LNPs) (Sahin et al., 2014) which can be tracked over time using automated microscope systems. Example LNP-based time-lapse data for our project is shown in **Figure 1**. In our modelling we are exploring to what extent successful drug delivery can be predicted in advance from temporal changes in cell morphology - via bright field and cell stain image channels (a counterstain and an LNPs/mRNA cargo stain).
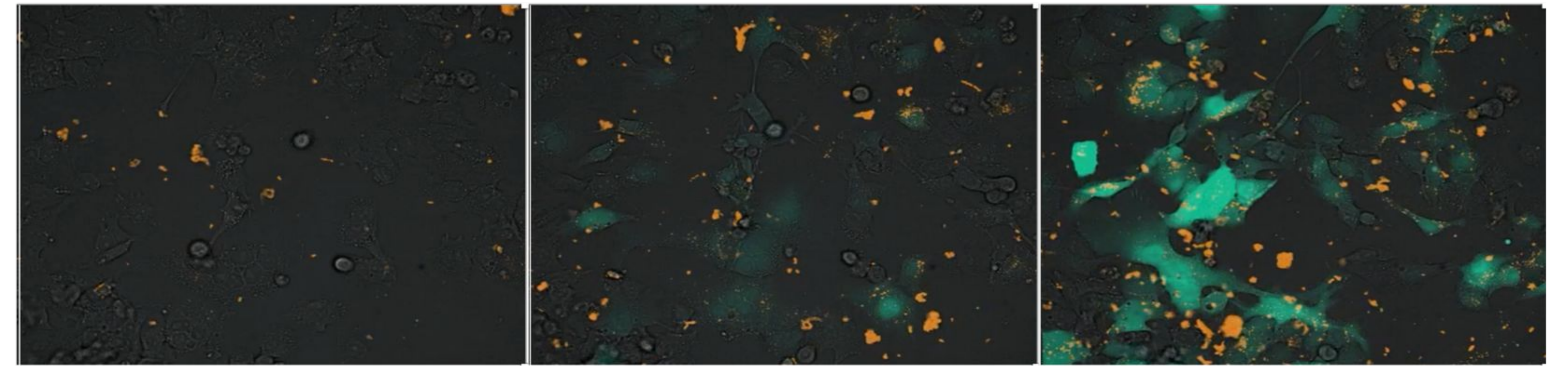
Figure 1: Frames from time-lapse data at the start (t=1), middle (t=36) and end (t=72) of an experiment for one well. The circular objects, detected by the bright field, show the cells; the orange parts show the LNPs; and the green the successfully expressed GFPs. Each image is 2554 x 2154 pixels.
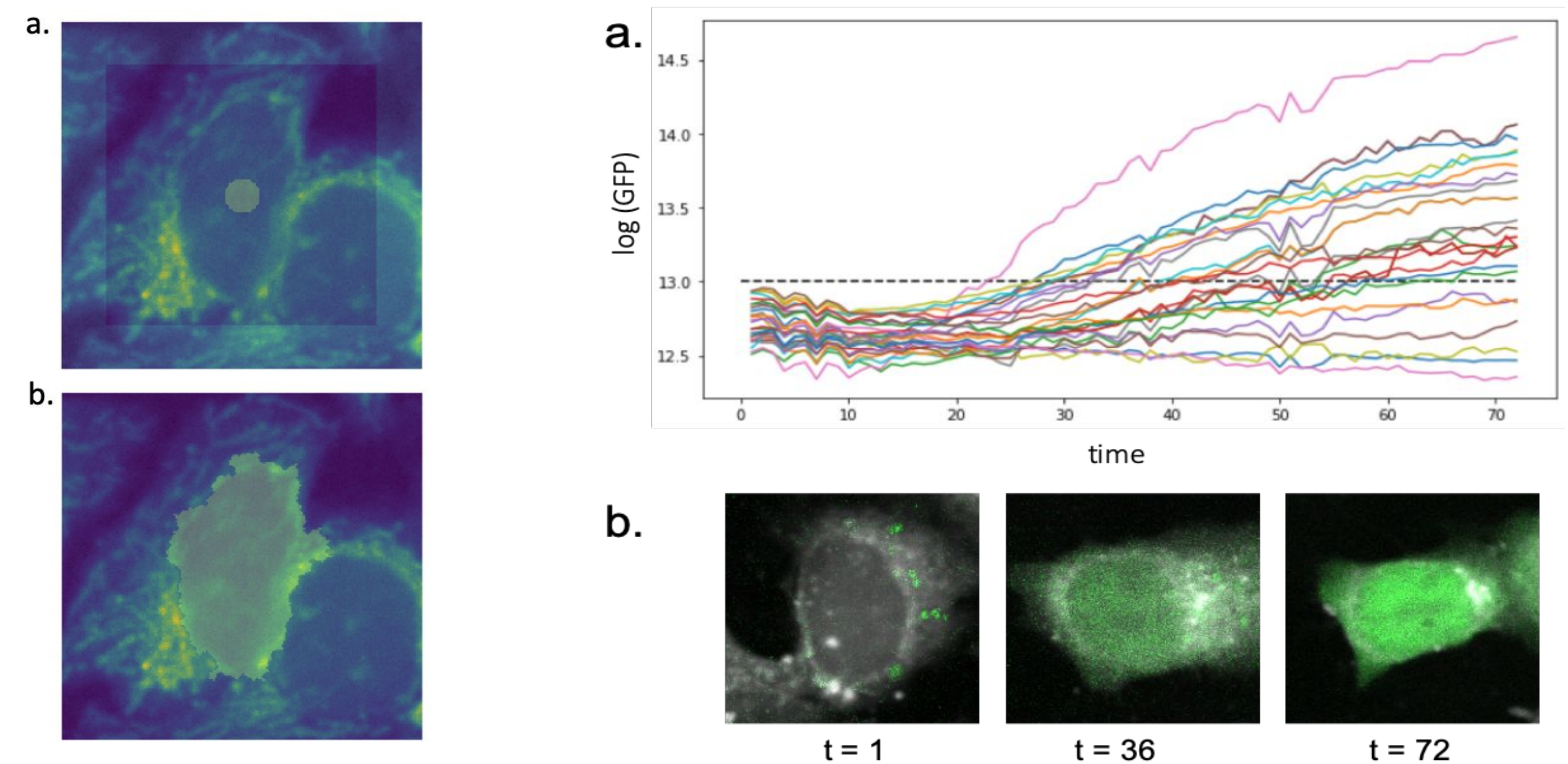
*Figure 2: **a)** the internal (seed) and external markers used and **b)** an example segmentation - using the counterstain channel - for a cell based on the seeded watershed algorithm.*

*Figure 3: **a)** GFP expression for cells within one well; **b)** snapshots of one cell at the start, middle and end of the experiment with high GFP expression. Dashed horizontal line in **(a)** shows the cutoff for successful GFP expression.*

Figure 4: Cell time-series modelling framework combining convolutional and recurrent neural networks. Inputs are 2D (multiple channeled) images which are passed through a set of convolutional and pooling operations. The representations are flattened and fed through a set of fully connected layers to encode feature vectors in each time point. These vectors are provided as input into a recurrent neural network (an LSTM) in which hidden states, h - which evolve through time - are trained to retain a memory of events relevant for the predictive purposes of the model. The output from the LSTM, after going through some final fully connected layers, is the predicted GFP expression at the final time point. Figure inspiration from Kimmel et al. (2019).

Modelling

Our modelling - at the cell-level at 40x magnification - is based on 192 x 192 pixel cut-outs around the center point of the cell at each time point. Nuclei center points are manually annotated in the first time point. As the cells do not move significantly between time points a seeded watershed segmentation (**Figure 2**) around the center point, followed by re-centering - providing the seed for the following time point - was sufficient to pre-process the data for the models. GFP expression through time for an example well is shown in **Figure 3**. As GFP expression does not begin until around t=22, we ask: can we predict the final GFP expression (at t=72) from the other imaging channels using only data from t=1 to t=22? The modelling framework, combining convolutional and recurrent neural networks, is shown in **Figure 4**.
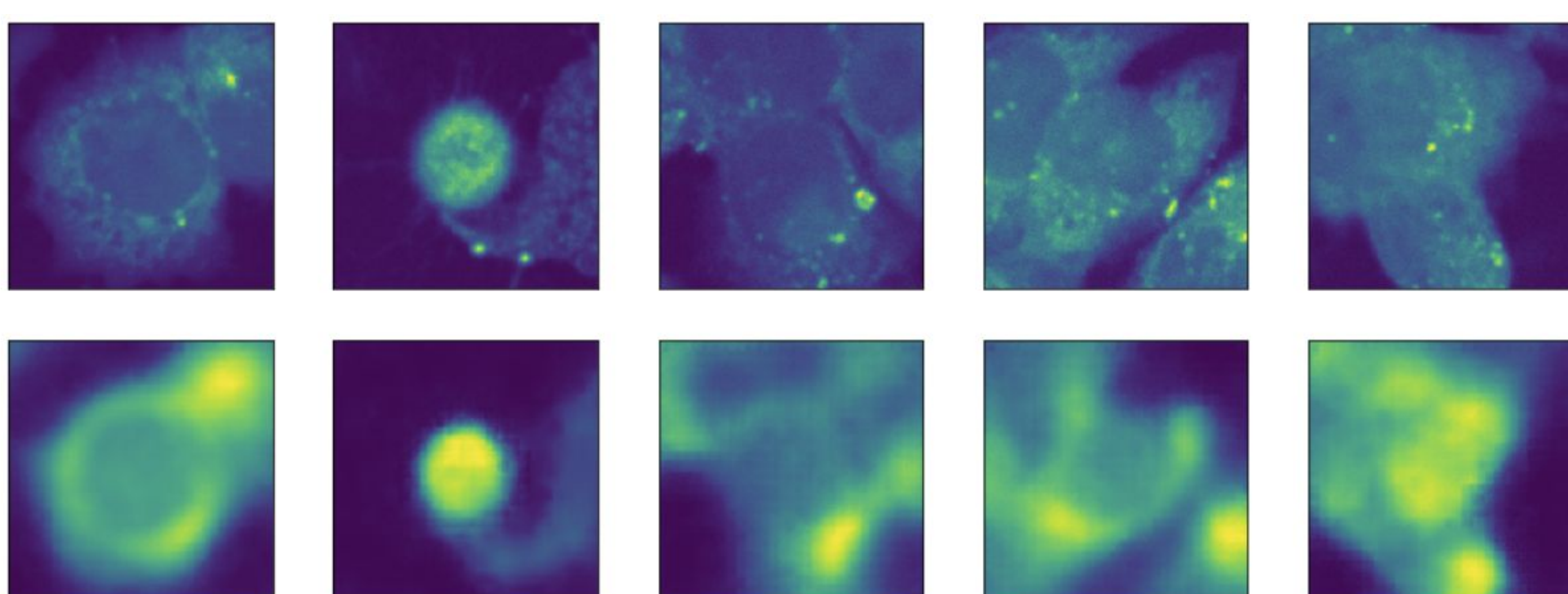
*Figure 5: Convolutional autoencoder reconstructions (bottom row) from a random selection of cell cutouts (192 x 192 pixels, counterstain images, top row) after going through a bottleneck vector of size 32 (producing the feature vectors as shown in **Figures 4** and **6**).*

Preliminary results
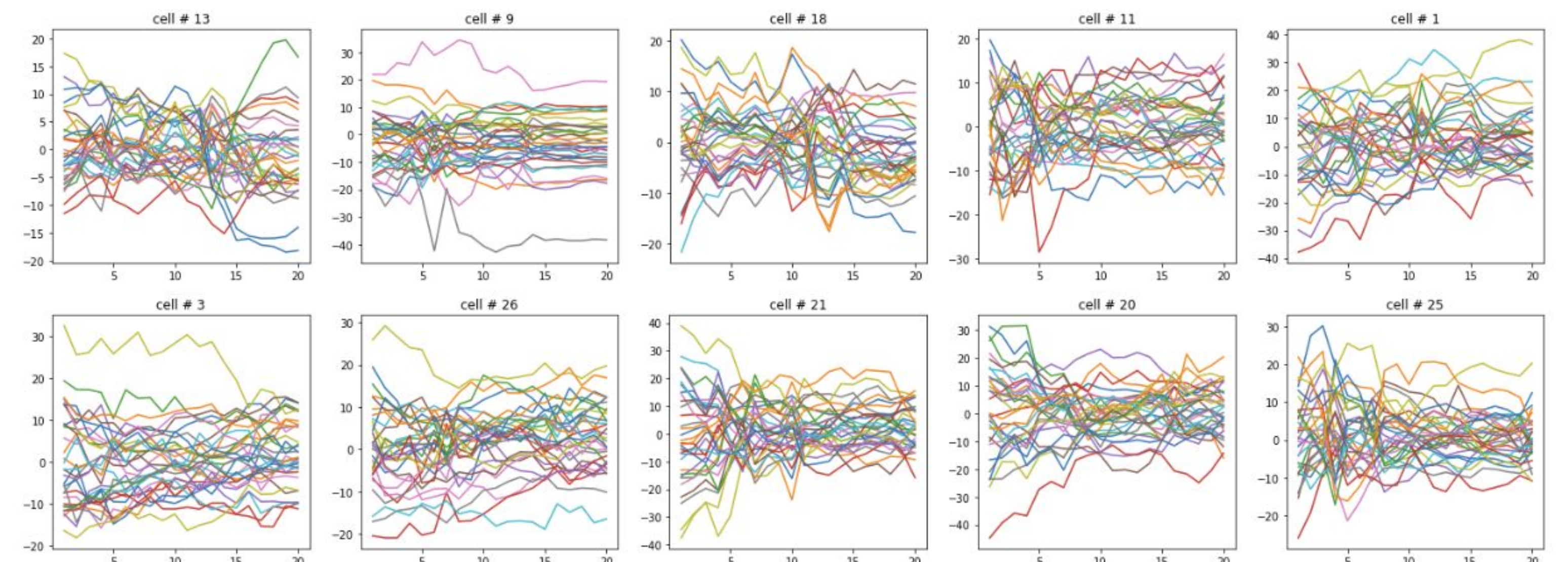
Using an initial exploratory dataset (3 x 3 wells), convolutional autoencoders for learning latent feature vectors (see **Figure 5**) and a bidirectional LSTM, we obtained significantly better GFP predictions using LNP input images as opposed to cell counterstain images (mean squared error (MSE) on validation data: 0.43 for LNP images and 0.89 for counterstain images). A selection of trajectories of the feature vectors fed into the LSTM are shown in **Figure 6**. It is difficult to see patterns in these trajectories by eye, hence the need for machine learning!

Next steps

We are presently collecting a larger dataset (16 x 16 wells), with a nuclear stain at t=1, to remove the need for an initial manual annotation, to explore:
- using various combinations of the three input channels;
- autoencoders reconstructing future images -> more relevant feature vectors?;
- unfreezing the encoder weights and training end-to-end;
- pre-trained CNNs for the encoder part of the network.

Also to investigate: models composed of more traditional (i.e. non-deep learning) components, and hybrid models with both deep learning and traditional elements. E.g., the feature vectors encoded via the encoder can be replaced with features measured via `CellProfiler` (Carpenter et al., 2006). As an alternative to using LSTMs, time-series features can be extracted via `tsfresh` (Christ et al., 2018). These features can then be fed into other machine learning algorithms (such as random forests) to make predictions. These more traditionally learnt features have the benefit of easier interpretability for uncovering the biological drivers behind the predictions.


*Figure 6: Trajectories of latent feature vectors (from t=1 to t=22, based on counterstain images) learnt from the convolutional autoencoder and fed into the LSTM in **Figure 4** to make predictions (of final GFP expression at t=72). The trajectories are for the cells in **Figure 3(a)**, with the top row for the five lowest GFP expressors and the bottom row for the five highest.*

References

Carpenter, A.E. et al. (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biology 7, R100.
Christ, M., et al.(2018). Time Series FeatuRe Extraction on basis of Scalable Hypothesis tests. Neurocomputing 307, 72–77.
Kimmel, J.,et al. (2019). Deep convolutional and recurrent neural networks for cell motility discrimination and prediction. BioRxiv 159202.
Sahin, U., et al.(2014). mRNA-based therapeutics — developing a new class of drugs. Nature Reviews Drug Discovery 13, 759–780.

pharm b.io   AstraZeneca   HASTE   SWEDISH FOUNDATION for STRATEGIC RESEARCH