ADeS: a deep learning based Apoptosis Detection System for live cell imaging.

Alain Pulfer1,2,6, Diego Ulisse Pizzagalli1,3,6, Paolo Armando Gagliardi4,6, Lucien Hinderling4, Paul Lopez3, Romaniya Zayats6, Pau Carrillo-Barberà3,9, Paola Antonello1,4, Miguel Palomino-Segura7, Alessandro Giusti8, Marcus Thelen1, Luca Maria Gambardella8, Thomas T. Murooka5, Olivier Pertz6, Rolf Krause3, Santiago Fernandez Gonzalez1,*

1 Institute for Research in Biomedicine, Faculty of Biomedical Sciences, USI - Switzerland
2 Department of Information Technology and Electrical Engineering, ETH Zurich - Switzerland
3 Euler Institute, USI - Switzerland
4 Institute of Cell Biology, University of Bern, Switzerland
5 University of Manitoba, Winnipeg – Canada
6 Instituto de Biotecnologia y Biomedicina (BioTecMed), Universitat de València - Spain.
7 Centro Nacional de Investigaciones Cardiovasculares - Spain
8 Dalle Molle Institute for Artificial Intelligence, IDSIA - Switzerland

& Contributed equally
* Corresponding author santiago.gonzalez@irb.usi.ch

Abstract

Live-cell imaging allows the study of apoptosis at cellular level, highlighting morphological hallmarks such as nuclear shrinkage, membrane blebbing, and cell disruption. Identifying the exact location and timing of this process is essential to foster the understanding of its spatial-temporal regulation. However, the analysis of live-cell imaging datasets is complex, whereas computational tools tailored to this task are yet scarce. Therefore, we developed ADeS, an Apoptosis Detection System based on deep learning and activity recognition. ADeS uses morpho-dynamic hallmarks to detect the exact location and timing of apoptotic events in different cell types, reaching an accuracy above 97% in the classification of our validation datasets acquired in vitro and in vivo. Moreover, ADeS is the first successful implementation of a deep learning network for the automatic detection of apoptotic cells in full microscopy movies in an end-to-end fashion, outperforming human in the same task. As a case study, we employed ADeS for the analysis of cell survival in vitro, and for tissue damage assessment in vivo, showing its potential application for toxicity assays, treatments evaluation and measuring of tissue dynamics.

Keywords: Apoptosis, Cell death, Intravital 2-photon microscopy, Bioimaging analysis, Artificial intelligence, Activity recognition, Computer vision, Spatial Biology.
Introduction

Programmed cell death refers to a biological process in which a cell ceases its functions to benefit the multi-cellular organism, e.g., in response to localized damage or to remove aged cells in favor of new ones. Amongst the different forms of programmed cell death, apoptosis is one of the best characterized. This process relies on the proteolytic activation of caspase-3-like effectors, which yields a series of morphological changes that include cell shrinkage, chromatin condensation, DNA fragmentation, and membrane blebbing. In its final stage, apoptosis culminates in the disruption of the cell into apoptotic bodies, which are vesicles that encapsulate the cytoplasm to prevent its leakage into the extracellular space. Apoptosis plays a crucial role in multiple biological functions, including tissue homeostasis, host protection, and immune response. Consequently, dysregulation of this process can lead to several pathological conditions, including inflammatory diseases and cancer.

Being able to determine the exact location and timing of apoptotic events is crucial to expand our understanding of the process. To identify apoptosis, classical imaging techniques rely on cellular staining on fixed cultures and tissues. However, this approach precludes the temporal visualization of the apoptotic process and introduces artifacts caused by the fixation of the sample. This problem can be addressed by live-cell imaging, which unravels the dynamic aspects and interplay of imaged cells. In vitro live-cell imaging offers several advantages, including tight control of the experimental conditions and compatibility with high-throughput microscopy. However, in vitro assays might not fully capture the complexity of the physiological environment.

As a complementary technique, intravital microscopy (IVM) enables live cell imaging in situ across different organs, becoming the most accurate model to describe cellular activities within a living host and allowing the study of apoptosis in physiological and pathological conditions. In particular, multi-photon intravital microscopy (MP-IVM) generates in depth 3D data that encompass multiple channels up to several hours of acquisition, providing unprecedented insights into cellular dynamics and interactions. Nonetheless, resulting data are complex and challenging to analyze, whereas automatized quantification approaches are scarce.

In addition, while apoptotic biomarkers are helpful for the quantification of apoptosis, their usage is not trivial and they can present technical limitations. For instance, caspase reporters can activate several minutes before a cell displays the apoptotic hallmarks. Meanwhile, these cells can at times revert the apoptotic process and regain a healthy phenotype. Consequently, biomarkers are suitable for a cumulative quantification of overall apoptosis, but they can face challenges in achieving high temporal accuracy. Tailored computational routines could address these limitations by automatically detecting individual apoptotic cells with high spatial and temporal accuracy. In this matter, approaches based on deep learning (DL) and activity recognition (AR) could improve the analysis of live-cell imaging data by detecting apoptotic events based on morphological patterns. However, to date, no DL method can process entire microscopy movies in an end-to-end fashion, namely by receiving a movie as an input and providing a spatial-temporal depiction of the apoptotic events as an output.
Therefore, we developed a novel Apoptosis Detection System (ADeS) designed to detect cell death events with high-resolution in time-lapses generated in vitro and in vivo. ADeS employs a recurrent DL architecture that computes the location and duration of multiple apoptotic events in live-cell imaging, representing a novel solution for apoptosis detection. Here, we show that ADeS is compatible with different imaging modalities, cellular staining, and cell types.

**Results**

Curated and high-quality datasets containing numerous instances of training samples are critical for the development of data-hungry methods such as supervised deep learning (DL) algorithms. To this end, we generated two distinct datasets encompassing epithelial cells (in vitro) and leukocytes (in vivo) undergoing apoptotic cell death. In addition, the two datasets include different imaging modalities (confocal and intravital 2-photon), biological models, and training-set dimensionalities. In this regard, a major difference between the datasets regards the staining methods and the morphological hallmarks that define the apoptotic process in both models. In the in vitro model, the expression of nuclear markers allowed us to observe apoptotic features such as chromatin condensation and nuclear shrinkage.

While in the in vivo model, cytoplasmic and membrane staining highlighted morphological changes such as membrane blebbing and the formation of apoptotic bodies. Accordingly, we have manually annotated these datasets based on the presence of the specific hallmarks, and each dataset includes two class labels depicting either apoptotic or non-apoptotic cells. The two datasets constitute the first step towards the creation, testing and validation of our proposed apoptosis detection routine.

**Generation of in vitro and in vivo live-cell imaging data.** To generate a dataset of apoptotic and non-apoptotic events in vitro, we used epithelial cells, which, among the human tissues, has the highest cellular turnover driven by apoptosis. This guarantees the removal of old or damaged cells, ensuring tissue homeostasis and preventing the onset of skin diseases such as cancer. From the bioimaging perspective, epithelia is a densely packed tissue with almost no extracellular matrix, which makes the analysis extremely challenging. Therefore, in epithelial research there is a concrete need for computational tools to identify apoptotic events automatically. To this purpose, we engineered the human mammary epithelial cell line MCF10A for the expression of a nuclear fluorescent marker (H2B-miRFP703; Fig.1A). From live-cell imaging movies in different experimental conditions, we manually annotated the centroids (x,y,t) of 13120 apoptotic nuclei. Moreover, we automatically segmented 301630 non-apoptotic nuclei (Fig. 1B, Supplementary 1A). The criteria of manual annotation were based on nuclear shrinkage and chromatin condensation, two of the most prototypical hallmarks of apoptosis (Fig. 1C). Successively, using the annotated coordinates of apoptotic and non-apoptotic nuclei, we cropped short time-lapse sequences with constant size and frame number (Fig. 1C). From the generated time-lapses, we confirmed that non-apoptotic nuclei had constant area and chromatin density, whereas apoptotic nuclei displayed a decrease of the area and an increase of the chromatin...
condensation (Fig. 1D). The resulting dataset captured the heterogeneity of apoptotic cells in an epithelial tissue, including early nuclear fragmentation, rapid shift along x and y axes and extrusion through the z dimension (Supplementary Fig. 1B-C). Moreover, our dataset incorporates the typical difficulties of automatically annotating apoptotic events from live microscopy of a densely packed tissue (Supplementary Fig. 1D), with accumulation of apoptotic bodies (Supplementary Fig. 1E) and across multiple microscope hardware settings (Supplementary Fig. 1F).

For the generation of an in vivo dataset, we focused on polymorphonucleated leukocytes (neutrophils and eosinophils) expressing GFP. In these early immune responders, apoptosis is a crucial process that orchestrates their disposal, consequently determining the duration of the inflammation. To acquire instances of apoptotic leukocytes, we performed multiphoton intravital acquisitions (MP-IVM) in anesthetized mice by surgically exposing either the spleen or the popliteal lymph node (Fig 1E-F). The resulting time-lapses (Fig.1G) were 3D imaging data encompassing consecutive multi-focal planes (3D) and multiple imaging channels. From the generated MP-IVM movies, we manually annotated the centroid of apoptotic cells (Fig. 1H) by tracking them over time. For each apoptotic event, we generated cropped sequences of fixed size that tracked apoptotic cells for the duration of their morphological changes (59x59 pixels + time; Fig. 2E). This procedure was applied to 30 MP-IVM movies, generating 120 apoptotic sequences (supplementary Fig. 1G). Furthermore, we annotated random instances of non-apoptotic events, generating 535 cropped samples. To characterize the heterogeneity of the movies, we manually quantified the cell number per the field of view (87 ± 76), the shortest distance between cells (21.2 μM ± 15.4), and the signal-to-noise ratio (8.9 ± 3.6; supplementary Fig. 1 H-J). For detection purposes, we assumed that the morphological changes associated with apoptosis occurs within defined time-windows. Hence, we estimated a median duration of the morphological changes corresponding to 8 frames (supplementary Fig. 2K-L, respectively). In addition, to classify apoptotic cells within defined spatial regions, we considered them as non-motile. This assumption was confirmed, as we found that apoptotic cells showed a displacement and speed that were not significantly different from those of arrested cells (supplementary Fig. 2M).

ADeS, a pipeline for apoptosis detection. Detection of apoptosis in live-cell imaging is a two-step task that involves the correct detection of apoptotic cells in the movies \((x,y)\), and the correct estimation of the apoptotic duration \((t)\). To fulfill these requirements, we designed ADeS as a set of independent modules assigned to distinct computational tasks (Fig. 2). As an input, ADeS receives a 2D representation of the microscopy acquisitions (Fig. 2A) that is obtained from the normalization of 2D raw data or from the maximum projection of 3D data. This processing step ensures the standardization of the input data, which might differ in bit depth or acquisition volume. Successively, ADeS employs a selective search algorithm to compute regions of interest (ROIs) that might contain apoptotic cells (Fig. 2B). For each ROI at time \(t\), ADeS extracts a temporal sequence of \(n\) frames ranging from \(t - n/2\) to \(t + n/2\) (Fig. 2C). The resulting ROI sequence is standardized in length and passed to a recurrent DL network (Supplementary Fig. 2A-B), which evaluates if it is apoptotic or non-apoptotic. Finally, each apoptotic sequence is depicted as a set of bounding boxes and associated probabilities (Fig. 2D).
generated from the predicted trajectories (x, y, t, ID; Fig. 2E). From this readout, ADeS can generate a heatmap representing the likelihood of apoptotic events over the course of a movie (Fig. 2F, left), altogether with a cumulative sum of the predicted cell deaths (Fig. 2F right).

ADeS training and deployment in vitro. As previously described, ADeS is a multiple-block pipeline, for which reason, its application and validation follow two main steps: 1) the training of the DL classifier with a target dataset, 2) its deployment on live-cell imaging acquisitions. As opposed to in vivo acquisitions, in vitro time-lapses are more homogeneous in their content and quality, representing thus the first dataset in order of complexity for the training ADeS. Therefore, we formulated the learning problem as a binary classification task in which non-apoptotic sequences were assigned to the class label 0, while apoptotic sequences were assigned to the class label 1 (Supplementary Fig. 3A). The class label 0 included instances of healthy nuclei and nuclei undergoing mitotic division, which can resemble apoptotic events. Moreover, we iteratively populated the class label 0 with false positives obtained from a preliminary DL network deployed on control videos without apoptotic events (Supplementary Fig. 3B), ensuring a systematic decrease of the misclassification rate.

Using the final training dataset, we trained the DL classifier for 100 epochs with a 0.12 validation split (Fig. 3A). After deploying the trained model on 1000 testing samples, the confusion matrix (Fig. 3B) displayed a very low misclassification rate (2.68%) that was similarly distributed between false positives (1.04%) and false negatives (1.64%). Accordingly, the receiver operating characteristic (ROC) of the model was skewed towards the left (AUC = 0.99, Fig. 3C), indicating a highly favorable tradeoff between true positive rate (TPR) and false positive rate (FPR), as already suggested by the overall predictive accuracy of 97.32% (Fig. 3B). Altogether, these results show an unprecedented accuracy of the DL model in the classification of apoptotic and non-apoptotic sequences. However, they reflect the theoretical performances of the classifier applied to cropped sequences depicting a single cell at a time. To validate ADeS on full-length microscopy acquisitions, we deployed it on six testing movies that were not part of the training set. Each testing movie has been annotated manually and it contains a variable number of ground truth apoptosis (98 ± 21) and a comparable cell density (1705 ± 124). Moreover, all movies had identical magnification (20x), duration (21 h), and sampling rate (5 min). For the testing of ADeS on these movies, we followed an unbiased approach, and we did not hard-tune the hyper-parameters of the model (see material and methods), specifying only a stringent confidence threshold (0.990) and a temporal window based on the average duration of the nuclear hallmarks (9 frames). As a result, ADeS could predict the location and timing of the apoptotic nuclei (Fig. 3D, Supplementary Movies 1-6), proving capable of detecting multiple apoptosis in densely packed field of view (Fig. 3E-F). To quantify these performances, we compared the prediction of ADeS to the annotated ground truths (x,y,t). Hence, we found that the average TPR, or sensitivity, was 78.5% (ranging from 68% to 92%), while the average FPR was 3.9%. The missed apoptotic events are likely due to the heterogeneity of nuclear fragmentation, which can vastly differ in signal intensity, size, focal plane and in duration (Supplementary Fig.1). Nonetheless, hard tuning the model could further increase the sensitivity without additional training data, for instance by adjusting the temporal interval or by lowering the confidence
threshold. Moreover, false positives appear to be contained, with a new false positive every 4 frames, or 20 minutes. This rate suggests that ADeS is very robust, especially considering a field of view depicting a high number of cells (more than 1700) at each frame.

With respect to the spatial-temporal dynamics, the apoptotic count over time highlights a tight relationship between ground truth apoptosis and correct detections of ADeS (Fig. 3H). Accordingly, the two curves are divergent but highly correlating (pearson r = 0.998), proving that ADeS can successfully capture cell death dynamics. The spatial-temporal proximity between ADeS and the ground truth is visually depicted by a 2D scatterplot \((x, y, t = \text{radius})\) indicating overlap between the two scatter populations (Fig. 3I). This relationship is further captured by nearest neighbor (NN) analysis, showing that the average distance between all ADeS predictions (true positives + false positives) and the nearest neighbor in the ground truth is 30 pixels, whereas randomly generated predictions would have a ground truth NN within a 52 pixels radius (Fig. 3J). Considering instead the true positives only, we observed that they are in close spatial proximity to the ground truth, with most predictions within a 20 pixel radius (Fig. 3K). The difference between the predicted timing of apoptosis and the one annotated in the ground truth is also small, with an average discard of 3.8 frames (Fig. 3L). It should be noted, however, that the temporal discard is less informative than the spatial one, as it is also representative of the bias of the operators that subjectively indicated the beginning of the apoptosis. Altogether, our quantifications indicate that ADeS detects apoptotic nuclei \textit{in vitro} with high spatial and temporal accuracy, establishing a novel comparative baseline for this task.

**ADeS training and deployment \textit{in vivo}.** Upon the successful application of ADeS \textit{in vitro}, the next step in complexity was the application to \textit{in vivo} time lapses, which are inherently more challenging due to high background signal, autofluorescence and the presence of collagen\textsuperscript{42}. For this purpose, we re-trained ADeS using the \textit{in vivo} data described in Figure 2. However, one of the main limitations of supervised ML is the need for large datasets. Therefore, the finite number of MP-IVM acquisitions and apoptotic instances represented a bottleneck for the training of ADeS. To overcome this limitation, we implemented a custom data augmentation strategy that exploits 3D volumetric rotations as previously performed in other studies \textsuperscript{43,44}. Accordingly, each 3D apoptotic sequence underwent multiple rotations and was successively projected in 2D (Supplementary Fig. 4A-B).

To train ADeS using the latter generated \textit{in vivo} dataset, we defined a binary classification task in which ROIs containing apoptotic cells were assigned to the class label 1, while any other ROIs, including healthy cells and elements of background, were assigned to the class label 0 (Supplementary Fig. 4C). Successively, we trained the DL classifier for 200 epochs performing 5-fold cross-validation according to the movies ID (Fig. 4A). The resulting confusion matrix showed a classification accuracy of 97.42\%, with a 2.58\% misclassification rate that was largely explained by a type II error (2.44\% false negatives) (Fig. 4B). Analogously to the tests \textit{in vitro}, classification \textit{in vivo} proved highly effective in predicting apoptotic and non-apoptotic instances. This favorable result was supported by the ROC of the model, which indicated high sensitivity and a moderate FPR (Fig. 4C).
To understand which features the model employed for its predictions in vivo, we classified multiple apoptotic sequences in which the cell membrane disrupted at shifted time points. Remarkably, we found that the network always attained the peak activation when the cell membrane collapsed (Fig. 4D). This observation was congruent with our dataset annotation strategy and confirmed the temporal awareness of the model. Successively, we benchmarked ADeS in the detection task performed on a set of 23 MP-IVM acquisitions of immune cells undergoing apoptosis. Differently from in vitro settings, in vivo acquisitions displayed high variability in cell number, auto-fluorescence, signal intensity and noise levels (Supplementary Fig. 4D). Regardless, ADeS correctly predicted the location and timing of cells undergoing apoptosis (Fig. 4H, Supplementary Movies 7-12), showing robustness to increasingly populated fields of view (Supplementary Fig. 4E). Moreover, the pipeline successfully applied to neutrophils imaged in the lymph node (Fig. 4E) and eosinophils imaged in the spleen (Fig. 4F). By comparing ADeS predictions with the annotated ground truths, we found that our pipeline detected apoptotic events with a TPR of 81.2% and a FPR of 4.6% (Fig. 4G). The detections, provided in the form of bounding boxes and trajectories, indicated the coordinates and duration of the events. Hence, to measure how close they were to the annotated trajectories, we employed the tracking accuracy metric (TRA), a compound measure that evaluates the similarities between predicted and ground truth trajectories. The average TRA was above 0.9, indicating high fidelity of the trajectories predicted by ADeS (Fig. 4H). Successively, we compared ADeS to human annotation performed by three operators on five testing movies. As a result, ADeS displayed a downward trend of the FPR, although there was no significant difference in the TPR and FPR (Fig. 4I). Nevertheless, ADeS performances appeared to distribute in two distinct groups: a large one, with an average sensitivity of 100%, and a smaller one, with an average sensitivity of 52% (Fig. 4J). To understand this discrepancy, we applied hierarchical clustering to the testing videos according to their imaging properties and biological content (Fig. 4J), generating two major dendrograms. The first dendrogram contained mostly videos with reduced sensitivity (yellow) and was defined by a high cell number, high noise levels, short cell distance, and a saturated and fluctuating image signal. Most prominently, the cell number played a crucial role in overall performance, as an increment of this parameter resulted in a pronounced decrease of the TPR and a moderate increase of the FPR (Fig. 4K). Incidentally, the positive predictive value (PPV) was significantly lower in videos with poor SNR and, although not statistically significant, the PPV was lower when the signal standard deviation was higher (Fig. 4L). Following these observations, we hypothesized that the overall quality of a movie is a predictor for ADeS performances. Hence, we combined the parameters highlighted by the clustering analysis (Fig. 4J) into a single score ranging from zero to one (one indicating the highest and ideal score), finding a weak correlation between the video quality and the sensitivity of ADeS (Fig. 4M). This trend was evident when we considered only videos with suboptimal sensitivity, in which case we found a strong correlation (0.72) confirming that the video quality partially explains the observed performances (Fig. 4N).

Finally, we evaluated how the biological variability in vivo could affect the readout of ADeS. For this purpose, we defined nine distinct biological categories that included instances of apoptotic cells, healthy
cells and background elements. For all biological categories, the classification accuracy was superior to 80%, except for cells overlapping and cells with high membrane plasticity (Supplementary Fig. 4F).

**Biological applications of ADeS in vitro.** A common application of cell death staining is the evaluation of the toxicity associated with different compounds\(^1\,\)\(^{13}\,\)\(^{45}\), or the efficacy of an apoptotic-inducing treatment. Here, we show that ADeS has analogous purposes and can effectively quantify the toxicity of different compounds in vitro. For this application, we grew epithelial cells in vitro, treating them with PBS and three increasing concentrations of doxorubicin, a chemotherapeutic drug that elicits apoptosis in the epithelium\(^46\). Epithelial cells were seeded with a density of 105 cells per well, and all four conditions had the same confluence before the treatment. However, at 24 h. post-acquisition, the number of survivor cells was inversely proportional to the doxorubicin concentration (Fig. 5A). We confirmed this trend with ADeS, which measured the lowest mortality after 24 h. in PBS (64 cells), followed by doxorubicin concentrations of 1.25 \(\mu\)M (95 cells), 2.50 \(\mu\)M (161 cells) and 5.00 \(\mu\)M (283 cells). Moreover, ADeS predicted distinct pharmacodynamics (Fig. 5B) which can define the drug concentration and experimental duration required to reach a specific effect in terms of apoptotic count.

To this end, each time point in figure 5B also defines a dose-response relationship. Here, we provide two dose-responses curves, at 5 h. and 24 h. post treatment, showing different pharmacodynamics (EC\(50\) 5h = 2.35, Hill slope 5h = 3.81, EC\(50\) 24h = 4.47, Hill slope 24h = 1.93, Fig. 5C-D). Notably, the fit can project the dose-responses for higher drug concentrations, predicting the maximum effect size at a given time. For instance, at 24 h. post treatment, a 10 \(\mu\)M titration attains 86\% of the maximum effect (456 apoptotic cells), whereas further increasing the concentration of the drug leads only to a slow reaching of the plateau (Fig. 5E). We suggest that this approach is beneficial to maximize the effect of a drug towards a designated target, while minimizing collateral damages on non-target cells. For instance, in chemotherapies employing doxorubicin, apoptosis of epithelial cells is an undesired effect. Therefore, researchers can select a titration of the drug and a duration of the treatment that does not affect the epithelium but is still being effective on the tumor. We finally demonstrate the reproducibility of the toxicity assay by targeting another cell type (T-cells) treated with a different apoptotic inducer (staurosporine, Supplementary Fig. 4).

**Biological applications of ADeS in vivo.** To test the application of ADeS in an in vivo setting, we evaluated the detection of apoptotic cells following laser ablation in the spleen of an anesthetized mouse (Fig. 6A). This method was previously used to study immune cell responses to tissue damage\(^47\). The insult caused prompt recruitment of neutrophils, which led to the formation of a local cluster (Fig. 6B, left, Supplementary Movie 13). Successively, the neutrophils within the swarm underwent apoptotic body formation in a coordinated manner, i.e. they disrupt simultaneously (Fig. 6B, right). To quantify this event, we processed the generated time-lapse with ADeS, resulting in a probability map of apoptotic events over the course of the acquisition (x,y,t,p) (Fig. 6C). Accordingly, the location with the highest probability corresponded to the area damaged by the laser, and the visual representation of the probability map enabled us to infer the morphology and location of the cluster. This result depicts the potential application of ADeS in digital pathology, showing how the distribution of apoptotic events
throughout the tissue can identify areas enriched by cell death events. In addition, it could be used to measure the degree of photo damage induced during the imaging process.

Finally, we applied ADeS to study the response of bystander cells following apoptotic events in the lymph node of mice treated with an influenza vaccine. We computed the spatial and temporal coordinates of a neutrophil undergoing apoptosis (Fig. 6D), which in combination to the tracks of neighboring cells allowed us to characterize patterns of cellular response that followed the apoptotic event. Amongst other parameters, we observed a sharp decrease of the distance of the neighboring cells with respect to the apoptotic centroid (Fig. 6E), altogether with a pronounced increase of the cells instantaneous speed (Fig. 6F). These observations, associated with cell polarization towards the site of cell death, are in agreement with studies suggesting that apoptosis is not a silent process.

Discussion

Automated bio-image analysis precludes the need for manual annotation and avoids bias introduced by the researcher. In this regard, DL and AR techniques led to successful applications in tissue classification and cell segmentation. Moreover, for the quantification of cell death, recent studies demonstrated the usage of DL networks to classify small images and time-lapses containing single apoptotic cells. However, this approach is not suitable for large microscopy acquisition, as it does not address two fundamental questions of activity recognition, namely the location where an event occurs and the time component. These questions define the detection task, which has a computational cost that can rapidly grow with the size and the duration of a movie. Moreover, live-cell imaging data present specific challenges which further increase the difficulty of developing detection routines, including densely packed fields of view, autofluorescence and imaging artifacts among others. Consequently, computational tools to effectively detect and quantify apoptotic events in live-cell imaging remains unavailable. Therefore, our objective was the creation of an apoptosis detection pipeline that could address the aforementioned challenges, in the detection of apoptosis in vitro and in vivo. ADeS represents an important bridge between AR and bioimaging analysis, as it is the first apoptosis detection routine applicable to full microscopy time-lapses, whereas previous methods were not directed towards this aim. In addition, to encourage reproducibility and to foster the development of apoptosis detection routines, we presented two datasets encompassing multiple cell types, fluorescent labels, and imaging techniques. With this resource, researchers will be able to design and benchmark novel solutions for cell death detection. Moreover, these datasets can establish the basis for comparative studies on cell behavior in the context of cell death.

In terms of performance, and in agreement with studies on kinetics datasets, we showed that recurrent DL networks are suitable for AR and, specifically, for the classification of apoptotic events. In this regard, we highlighted the importance of benchmarking ADeS on a validation set that included challenges related with live-cell imaging, such as overlapping cells and low signal-to-noise samples, among others. Consequently, in the classification task ADeS achieved an accuracy that outperformed analogue studies. In the detection task on full microscopy movies, ADeS established instead a
novel standard, with performances comparable to imaging experts trained to identify apoptotic features. The latter result is especially relevant if we consider that, at present, human annotation constitutes the gold standard for the labelling of morphological changes used in supervised DL. However, human labor is time-consuming and affected by variability, whereas automatized routines are typically more robust and time efficient. Therefore, ADeS can find applications in the processing of high-throughput live-cell imaging, minimizing annotation time and research bias.

From a biological perspective, ADeS has multiple applications in fundamental and clinical research. Among others, ADeS can provide insights on pivotal cell death mechanisms, or it could monitor the therapies modulating apoptosis in various diseases as well as the toxicity of different compounds. In this regard, ADeS readout is analogue to standard fluorescent probes for apoptosis detection, with the advantage that it applies directly to nuclear or cytoplasmic staining. Therefore, ADeS does not employ any additional acquisition channel, which in turn can be used for multiplexing purposes. Moreover, common probes\cite{11-15} flag early stages of apoptosis and are activated up to several minutes before the morphological changes of the cell\cite{23,25} are elicited, meanwhile, these cells can reverse the apoptotic process\cite{26,27,53}. By contrast, ADeS indicates the exact instant of cell disruption, adding specificity to the temporal dimension. For these reasons, we suggest that ADeS can complement the information provided by classic apoptotic biomarkers, proving useful in experimental assays where the temporal resolution delivers more information than the sole apoptotic count. For instance, in the study of tissue dynamics, where the activity of single-cells can reveal connections between the underlying signaling pathways and the fate decision of individual cells, such as mitosis or apoptosis\cite{54,55}. A prototypical example reported in in vitro and in vivo epithelial systems is the extracellular signal-regulated kinase (ERK) wave triggered by single apoptotic events, a signal that expands radially to neighboring cells and prevents them to further undergo apoptosis, allowing the epithelial tissue to withstand strong apoptotic stimuli while maintaining its integrity\cite{56}.

Analogously, in vivo studies showed that cell death stimuli are spatiotemporal triggers modulating the activity of immune cells. For instance, we illustrated how cell death events can anticipate a cellular recruitment, confirming the key role of apoptosis as a non-silent process\cite{48}. In turn, cell recruitment in response to cell death can be pro-inflammatory, as observed in neutrophil swarms\cite{57} or post macrophages death\cite{58}, or it can be anti-inflammatory, as a physiological process mediated by recruited phagocytes\cite{48}. Such systems exhibit complex dynamics for which interpretation ADeS provides a comprehensive readout in terms of spatial and temporal coordinates. We suggest that the integration of these features, altogether with experimental observations, could provide the basis to unravel the mechanism of complex signaling pathways. Future work in this direction might include predictive models that link specific cell death dynamics to underlying stimuli and diseases associated with apoptosis.

In conclusion, ADeS constitutes a novel solution for apoptosis detection that combines state-of-the-art microscopy and DL. Its successful implementation represents a step towards the general application of AR methods to live-cell imaging, bridging two distinct fields and importing the advantages related to automated routines. Accordingly, ADeS could process large microscopy datasets in a robust and
unsupervised manner, finding relevant applications in fundamental cell death research and depicting its role in different tissues, diseases, and biological processes. Further work might extend the proposed pipeline to other cell populations, other types of cell death and, potentially, diverse cell actions.

Material and Methods

MCF10A cell line and image acquisition. The normal-like mammary epithelial MCF10A cells (provided by Joan Brugge69), stably expressing the nuclear marker, were generated as previously described56. Briefly, the nuclear marker H2B-miRFP703, provided by Vladislav Verkhusha (Addgene plasmid # 80001)60, was subcloned in the PiggyBac plasmid pPBbSr2-MCS. After cotransfection with the transposase plasmid61, cells were selected with 5 µg/ml Blasticidin and subcloned. For time-lapse imaging, the cells were seeded on 5 µg/ml fibronectin (PanReac AppliChem) coated 1.5 glass-bottom 24 well plates (Cellvis) at 1 x 10^5 cells/well density. After 48 hours, when the optical density was reached, the confluent cell monolayer was acquired every 1 or 5 minutes for several hours with a Nikon Eclipse Ti inverted epifluorescence microscope with 640nm LED light source, ET705/72m emission filter and a Plan Apo air 20x (NA 0.8) or a Plan Apo air 40x (NA 0.9) objectives. The collection of biological experiments used in this study includes different stimulation of apoptosis, such as growth factors and serum starvation and Doxorubicin, a chemotherapeutic drug, at various concentrations.

Apoptosis induction of MCF10A cells with doxorubicin. Normal-like mammary epithelial MCF10A cells were grown in 24 well glass coated with fibronectin with a seeding of 1x10^5 cells/well. After two days, cells were starved for three hours and treated with doxorubicin at 1.25, 2.50, and 5.00 μM concentrations.

Mice. Prior to imaging, mice were anesthetized with a cocktail of Ketamine (100 mg/Kg) and Xylazine (10 mg/Kg) as previously described 19. All animals were maintained in specific pathogen-free facilities at the Institute for Research in Biomedicine (Bellinzona, CH). All the experiments were performed according to the regulations of the local authorities and approved by the Swiss Federal Veterinary Office.

Intravital Two-Photon Microscopy. Surgery in the popliteal lymph node was performed as previously reported62. The exposed organs were imaged on a custom up-right two-photon microscope (TrimScope, LaVision BioTec). Probe excitation and tissue second-harmonic generation (SHG) were achieved with two Ti:sapphire lasers (Chamaleon Ultra I, Chamaleon Ultra II, Coherent) and an optical oscillator that emits in the 1,010–1,340 nm range (Chamaleon Compact OPO, Coherent) and has an output wavelength between 690–1,080 nm.

Neutrophil isolation from mouse bone marrow. Bone marrow samples were extracted via flushing with PBS from the long bones of UBC-GFP mice (https://www.jax.org/strain/004353). Then, the bone marrow was filtered through a 40um strainer and resuspended in PBS. Primary bone marrow neutrophils were isolated with Ficoll gradient and resuspended in PBS.
T-cell culture in a 3D collagen matrix. Human CD4+ T cells were isolated from the PBMC fraction of healthy donors obtained from NetCAD (Canadian Blood Services). Cell purity was above 95%. Naïve CD4+ T cells were activated by adding Dynabeads coated with anti-human CD3e/CD28 antibody (1:1 bead:cell ratio, Life Technologies Cat #11131D) in RPMI1640 supplemented with 10% FBS (VWR Seradigm Cat #1500-500), 2 mM GlutaMAX (Gibco Cat #3050-061), 1mM sodium pyruvate (Corning Cat #25-000-Cl) and 10mM HEPES (Sigma-Aldrich Cat #H4034). After two days, beads were removed and cells were cultured for another 4-6 days in a medium containing 50 IU/mL human rIL-2 (Biotechne Cat #202-IL-500), keeping cell density at 2 x 10^5 cells/mL. Cells were used for all experiments between days 6 to 8. All work with human blood has been approved by the University of Manitoba Biomedical Research Ethics Board (BREB).

Apoptosis live-cell imaging of T-cells in 3D collagen chambers. T-cells were labeled at day 6-8 using CMAC (10µM) cell tracker dye (Invitrogen) and glass slide chambers were constructed as previously described. Briefly, 2x10^6 cells were mixed in 270µL of bovine collagen (Advanced Biomatrix cat #5005-100ML) at a final concentration of 1.7 mg/mL. Collagen chambers were solidified for 45 minutes at 37°C/5% CO2 and placed onto a custom-made heating platform attached to a temperature control apparatus (Werner Instruments). For the induction of apoptosis, 1µM of Staurosporine (Sigma Cat #569397-100UG) and 800ng of TNF-a (Biolegend Cat #570104) in 100µL RPMI were added on top of the solidified collagen. Cells were imaged as soon as the addition of apoptosis inducers using a multiphoton microscope with a Ti:sapphire laser (Coherent), tuned to 800 nm for optimized excitation of CMAC. Stacks of 13 optical sections (512 x 512 pixels) with 4 mm z-spacing were acquired every 15 seconds to provide imaging volumes of 44mm in depth (with a total time of 60-120 minutes). Emitted light was detected through 460/50nm, 525/70 nm, and 595/50 nm dichroic filters with non-descanned detectors. All images were acquired using the 20X 1.0 N.A. Olympus objective lens (XLUMPLFLN; 2.0mm WD).

Data Processing and Image Analysis. The raw video data, composed by uint8 or uint16 TIFFs, were stored as HDF5 files. No video pre-processing was applied to the raw data before image analysis. Cell detection, tracking, and volumetric reconstruction of microscopy videos were performed using Imaris (Oxford Instruments, v9.7.2). The resulting data were further analyzed with custom Matlab and Python scripts (see code availability section).

Apoptosis annotation of epithelial MCF10A cells in vitro. We manually annotated apoptotic events of MCF10A cells by visual inspection of the movies. The annotation was done by observing the morphological changes associated with apoptosis (e.g. nuclear shrinkage, chromatin condensation, epithelial extrusion, nuclear fragmentation) across multiple consecutive frames. Using a custom Fiji macro, we automatically stored x and y centroids of the apoptotic nucleus. The time t of each apoptotic annotation was defined as the beginning of nuclear shrinkage.
Generation of the in vitro training dataset. The 16-bit raw movies were min-max scaled to the 0.001 and 0.999 quantiles and downsampling to 8-bit resolution. Using the database of manually labeled [x,y,t] coordinates of apoptotic events, we extracted crops with 59x59 pixels resolution (2x scaling for the FOVs acquired with the 20x objective). Seven time-steps of the same location were extracted, with linear spacing from -10 minutes to +50 minutes relative to the apoptosis annotation. This time frame was chosen to capture the cell before the onset of apoptosis, and the morphological changes associated with apoptosis (nuclear shrinkage, decay into apoptotic bodies, extrusion from epithelium). The resulting image cube has [x,y,t] dimensions of [59,59,7]. To create the training data for the non-apoptotic class, we excluded areas with an annotated apoptotic event with a safety margin from the movies. From the remaining regions without apoptoses, we extracted image cubes from cells detected with StarDist and from random locations. The random crops also included debris, apoptotic bodies from earlier apoptotic events, empty regions, and out-of-focus nuclei.

Apoptosis annotation of leukocyte cells in vivo. Three operators independently annotated the videos based on selected morphological criteria. To label apoptotic cells, the annotators considered only the sequences of cells that displayed membrane blebbing followed by apoptotic bodies formation and cell disruption (Fig 2 B). For each frame in the apoptotic sequence, the operators placed a centroid at the center of the cell with the Imaris “Spots” function, generating an apoptotic track. Successively, ground truth tracks were generated according to a majority voting system, and 3D volume reconstruction was performed on ground truth cells using the Imaris “Surface” function. Nearby non-apoptotic cells were also tracked. In addition, other non-apoptotic events were automatically subsampled from regions without apoptotic cells.

3D rotation of the in vivo annotations. In vivo annotations presented a class unbalance in favor of non-apoptotic cells, with a relative few apoptotic instances. Hence, to compensate for this bias, we produced several representations of the raw data by interpolating the raw image stacks in 3D volumes and rotating them in randomly sampled directions, with rotational degrees between 0° and 45°. After each manipulation, the rotated volume underwent flattening by maximum projection and symmetric padding to preserve the original dimension. The 2D images were successively resized and cropped to match the 59x59 pixels input of the classifier. Finally, the training sequences were saved as uint8 grayscale TIFFs files.

Generation of the in vitro and in vivo training datasets. To detect apoptotic cells in microscopy acquisitions, we defined a 2D binary classification task in which apoptotic events are labeled with class 1, while non-apoptotic events belonged to the class label 0. The resulting unprocessed data consisted of frame sequences composed of 3D crops. The content of the class label 0 in vitro included: healthy nuclei, background, cell debris and mitotic cells. The content of the class label 0 in vivo included: motile cells, arrested cells, highly deformed cells, overlapping cells, cell debris or blebs, empty background, noisy background, and collagen.
**Data augmentation and data loader.** Given the varying length of the training sequences contained in the TIFFs, upon training, we used a custom data loader that uniformly samples the input data and produces sequences with a fixed number of frames. The fixed number of frames was set to 5, corresponding to the frame-length of the shortest apoptotic sequence. During training, each sample underwent horizontal shift, vertical shift, zoom magnification, rotation, and flipping. All data augmentations were performed in python using the Keras library.

**Deep learning architecture.** As a deep learning classifier, we employed a recurrent architecture relying on time-distributed convolutional layers stacked on top of a long short-term memory layers (ConvLSTM). The input size consists of 5 single-channel images and pixel size of 59x59. The convolutional network has three layers of size 64, 128, and 256 length. Each layer has a 3x3 kernel, followed by Relu activation, batch normalization, and a dropout set to 0.3. The inclusion of padding preserves the dimension of the input, while 2D max-pooling is at the end of each convolutional block. After 2D max pooling, the output is passed to a unidirectional LSTM with 64 nodes, and successively to a fully connected decision layer. The fully connected network has four layers with 1024, 512, 128, and 64 nodes, each one followed by Relu activation and a 0.3 dropout layer. The last layer is a softmax activation, which predicts a decision between the two classes. The final model has 1,892,738 trainable parameters and 896 non-trainable parameters.

**Training and hyper-parameters.** Our model was trained in TensorFlow with Adam optimizer, using binary cross-entropy loss and an initial learning rate of 0.0001. The optimal mini-batch size was 32, and the number of training epochs was 200. In training mode, we set a checkpoint to save the model with the best accuracy on the validation dataset, and a checkpoint for early stopping with patience set to 15 epochs. In addition, the learning rate decreased when attending a plateau.

**Region proposal.** For the deployment of the classifier on microscopy videos, we applied the selective search implementation of OpenCV, generating a set of ROIs for each candidate frame of the input movie.

**Classification.** For each ROI computed by the region proposal at time t, a temporal sequence is cropped around t and classified with the Conv-LSTM. The resulting bounding boxes are filtered according to a probability threshold and processed with the non-maxima suppression utils from Pytorch.

**Centroid tracking.** Consecutive bounding boxes classified as apoptotic are tracked with a custom multi-object tracker based on Euclidean distance. The generated trajectories are filtered by discarding tracks with less than two objects.

**Default and user-defined parameters.** ROIs detected with the region proposal are filtered according to their size, discarding the ones with edges below 20 pixels and above 40 pixels. Furthermore,
threshold on intensity is applied to exclude uint8 patches with an average brightness below 40. Upon classification, a temporal window corresponding to the expected duration of the apoptotic event is set by the user (9 frames by default). This temporal window is subsampled to match the number of input frame of the classifier (5). The filtering of the predictions depends on a user-specified threshold, which by default corresponds to 0.95 in vivo and 0.995 in vitro. Non-maxima suppression is based on the overlapping area between bounding boxes, set to 0.1 by default. The centroid tracking has the following adjustable parameters: gap and distance threshold. The "gap" parameter, set to three frames, specifies for how long a centroid can disappear without being attributed a new ID upon reappearance. A threshold on the distance, set by default to 10 pixels, allows the connection of centroids within the specified radius.

All the reported quantifications had default parameters.

**Statistical analyses.** Statistical comparisons and plotting were performed using GraphPad Prism 8 (Graphpad, La Jolla, USA). All statistical tests were performed using non-parametric Kruskal-Wallis test or Mann-Witney test For significance, p value is represented as * when p < 0.05, ** when p < 0.005 and *** when p < 0.0005.

**Code availability**
The source code is available under the GPL v3 licence in [http://www.gitbub.com/IRB-LTDB/apoptosis-](http://www.gitbub.com/IRB-LTDB/apoptosis-) detection including the modules to

- extract the training set from in vivo data (Matlab)
- extract the training set from in vitro data (Matlab)
- 3D augmentation (Matlab)
- data loader (Python)
- generate region proposals (Python)
- train the Conv-LSTM model (Python)
- track ROIs (Python) and post-process results (Matlab)
Figure 1: Generation of in vitro and in vivo live-cell imaging data. A. Micrographs depicting mammary epithelial MCF10A cells transduced with H2B-miRFP703 marker and grown to form a confluent monolayer. The monolayer was acquired with a fluorescence microscope for several hours with 1-, 2-, or 5-min time resolution. B. The centroid (x, y) and the time (t) of apoptotic events were annotated manually based on morphological features associated with apoptosis. Non-apoptotic cells were identified by automatic segmentation of nuclei. C. Image time-lapses showing a prototypical apoptotic event (upper panels), with nuclear shrinkage and chromatin condensation, and a non-apoptotic event (bottom panels). D. Charts showing the quantification of nuclear size (left) and the standard deviation of the nuclear pixel intensity (right) of apoptotic and non-apoptotic cells (n = 50). Central darker lines represent the mean, gray shades bordered by light colored lines represent the standard deviation. E. Simplified drawing showing the surgical set-up for lymph node and spleen. F-G. Organs are subsequently imaged.
with intravital 2-photon microscopy (IV-2PM, F), generating 3D time-lapses (G). H. Representative IV-2PM micrograph and I. selected crops showing GFP-expressing neutrophils (white) undergoing apoptosis. The apoptosis sequence is depicted by raw intensity signal (upper panels) and 3D surface reconstruction (bottom panels).

**Figure 2. Pipeline for the activity recognition of apoptosis (ADeS).** A. ADeS input consists of single channel 2D microscopy videos (x,y,t). B. Each video frame is pre-processed to compute the candidate Regions of Interest (ROI) with a selective search algorithm. C. Given the coordinates of the ROI at time t, ADeS extracts a series of snapshots ranging from t-n to t+n. The convoluted long-short term memory (Conv-LSTM) classifies the sequence either as non-apoptotic (0) or apoptotic (1). D. The predicted apoptotic events are labelled at each frame by a set of bounding boxes which, E. are successively linked in time with a tracking algorithm based on Euclidean distance. F. The readout of ADeS consist of bounding boxes and associated probabilities, which can generate a probability map of apoptotic events over the course of the video (left) as well as providing the number of apoptotic events over time (right).
Figure 3. ADDeS training and performance in vitro. A. Confusion matrix of the trained model at a decision-making threshold of 0.5. B. Receiver operating characteristic (ROC) displaying the false positive rate (FPR) corresponding to each true positive rate (TPR). C. Training accuracy of the final model after 100 epochs of training. D. Representative example of apoptosis detection in a time-lapse in vitro acquisition. E. Multiple detection of nuclei undergoing apoptosis displays high sensitivity in densely packed field of views. F. Heatmap representation depicting all apoptotic events in a movie and the respective probabilities. G. Bar plots showing the true positive rate (TPR) and false positive rate (FPR) of ADDeS applied to five testing movies. H. Time course showing the cumulative sum of ground truth apoptosis (blue) and correct predictions (red). I. 2D visualization of spatial-temporal coordinates of ground truth (blue) and predicted apoptosis (red). In the 2D representation, the radius of the circles maps the temporal coordinates of the event. J. Pixel distance between ADDeS predictions and the nearest neighbor (NN) of the ground truth (left) in comparison with the NN distance obtained from a random distribution (right). The plot depicts all predictions of ADDeS, including true positives and false positives. K. Scatterplot of the spatial distance
between ground truth and true positives of ADeS. Ground truth points are centered on the $X = 0$ and $Y = 0$
coordinates. Red dashed circle indicates a 20 pixel radius around the ground truth origin. L. Temporal distance (frames) between the correct predictions of ADeS and the annotated time points in the ground truth.

Figure 4. ADeS training and performance in vivo. A. Confusion matrix of the trained model at a decision-making threshold of 0.5. B. Receiver operating characteristic (ROC) displaying the false positive rate (FPR) corresponding to each true positive rate (TPR). C. Training accuracy of the final model trained for 200 epochs with data augmentations. D. Image gallery showing ADeS classification to sequences with different disruption timing. The generated heatmap reaches peak activation (red) at the instant of cell disruption E. Representative snapshots of a neutrophil undergoing apoptosis. Green bounding boxes represents ADeS detection at the moment of cell disruption F. Representative micrograph depicting the detection of two eosinophil undergoing cell death in the spleen (left) and the respective probability heatmap (right). G. ADeS performances expressed by means of true-positive rate (TPR) and false-positive rate (FPR) over a panel of 23 videos. H. TRA measure distribution of the trajectories predicted by ADeS with respect to the annotated ground truth ($n = 8$) I. Comparison between human
and ADeS by means of TPR and FPR on a panel of 5 randomly sampled videos. Hierarchical clustering of several video parameters producing two main dendrograms (n = 23). The first dendrogram includes videos with reduced sensitivity and is enriched for several parameters related to cell density and signal intensity. K. Graph showing the effect of cell density on the performances expressed in terms of TPR and FPR (n = 13). L. Comparison of the positive predictive value of between videos with high and low signal to noise ratio (left), and videos with low and high shortest cell distance (right). M-N. Selected video parameters are combined into a quality score that weakly correlates with the TPR in overall data (M, n = 23) and strongly correlates with the TPR in underperforming conditions (N, n = 8). Statistical comparison was performed with Mann-Whitney test. Columns and error bars represent the mean and standard deviation respectively. Statistical significance is expressed as: p ≤ 0.05 (*), p ≤ 0.01 (**), p ≤ 0.001 (***), p ≤ 0.0001 (****).

Figure 5. Biological application of ADeS in vitro. A. Representative snapshots depicting epithelial cells in vitro at 0 and 24 hours after the addition of PBS and three increasing doses of doxorubicin, a chemotherapeutic drug and apoptotic inducer. B. Plot showing the number of apoptotic cells detected by ADeS over time for each experimental condition. C-D. Dose-response curves generated from the drug concentrations and the respective apoptotic counts at 5 and 24 hours post treatment. Vertical dashed lines indicates the EC50 concentration. E. Dose-response curve projected from the fit obtained in (D). The predicted curve allows to estimate the response at higher drug concentrations than the tested ones.
Figure 6. Biological application of ADeS in vivo. A. Schematic drawing showing the intravital surgical set up of a murine spleen after inducing a local laser ablation. B. Intravital 2-photon micrographs showing the recruitment of GFP-expressing neutrophils (Green) and the formation of a neutrophil cluster (red arrows) at 60 min after photoburning induction. C. Application of ADeS to the generation of a spatiotemporal heatmap indicating the probability of encountering apoptotic events in the region affected by the laser damage. The dashed circle indicates a hot spot of apoptotic events. D. Schematic drawing showing the intravital surgical set up of a murine popliteal LN after administration of an influenza vaccine in the mouse footpad. E. Intravital 2-photon micrographs showing ADeS detection of an apoptotic neutrophil (Blue, left) and the subsequent recruitment of neighboring cells (right) in the popliteal LN at 19 h following influenza vaccination. F. Plot showing the distance of recruited neutrophils with respect to the apoptotic coordinates over time (n = 22). G. Plot showing the speed of recruited neutrophils over time (n = 22). The dashed vertical line indicates the instant in which the apoptotic event occurs.

Bibliography


