# Weakly-supervised cell classification for effective High Content Screening*

Adriana Borowa[1,2][0000−0003−3387−6079], Szczepan Kruczek[3], Jacek Tabor[1][0000−0001−6652−7727], and Bartosz Zieliǹski[1,2][0000−0002−3063−3621]

[1] Faculty of Mathematics and Computer Science, Jagiellonian University, Łojasiewicza 6, 30-348 Kraków, Poland
ada.borowa@student.uj.edu.pl
{jacek.tabor, bartosz.zielinski}@uj.edu.pl
[2] Ardigen SA, Podole 76, 30-394 Kraków, Poland
[3] Selvita SA, Bobrzyńskiego 14, 30-348 Krakow, Poland
szczepan.kruczek@selvita.com

**Abstract.** High Content Screening (HCS) allows for a complex cell analysis by combining fluorescent microscopy with the capability to automatically create a large number of images. Example of such cell analysis is examination of cell morphology under influence of a compound. Nevertheless, classical approaches bring the need for manual labeling of cell examples in order to train a machine learning model. Such methods are time- and resource-consuming. To accelerate the analysis of HCS data, we propose a new self-supervised model for cell classification: Self-Supervised Multiple Instance Learning (SSMIL). Our model merges Contrastive Learning with Multiple Instance Learning to analyze images with weak labels. We test SSMIL using our own dataset of microglia cells that present different morphology due to compound-induced inflammation. Representation provided by our model obtains results comparable to supervised methods proving feasibility of the method and opening the path for future experiments using both HCS and other types of medical images.

**Keywords:** High Content Screening · Weakly-supervised learning · Self-supervised learning · Multiple Instance Learning.

## 1 Introduction

High Content Screening (HCS) is used to identify changes in a cell phenotype caused by e.g. small molecules or RNA [12]. The main application of HCS is the

* The works are carried out under contract no. POIR.01.01.01-00-0878/19-00, as: ,,HiScAI - Development of cell-based phenotypic platform based on high content imaging system integrated with artificial intelligence data analysis for neuroinflammatory and fibrosis drug discovery", co-financed by the European Regional Development Fund under the Smart Growth Operational Programme, Submeasure 1.1.1.: Industrial research and development work implemented by enterprises.

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_23

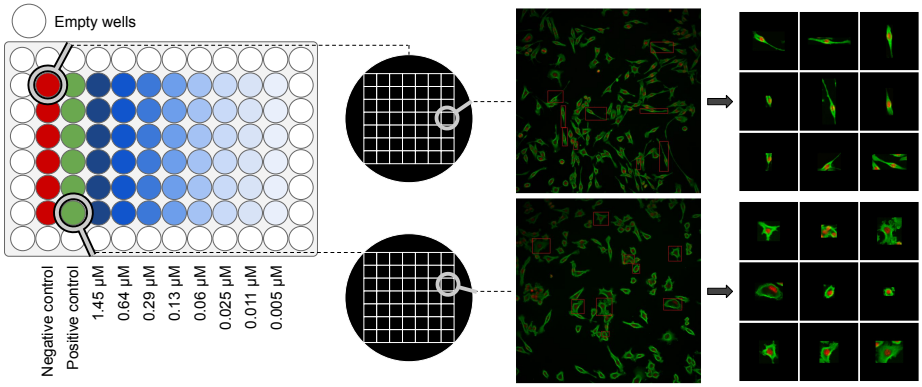


**Fig. 1.** High Content Screening data: Starting from left, the first part of the figure present an entire HCS plate containing 96 wells. Wells marked red contain negative control samples – without any compounds – cells are not inflamed and they are elongated with rather smooth borders and visible long processes. Wells marked green contain positive control samples with LPS induced inflammation cells, they get smaller, rounder, and develop short 'spikes'. Wells marked blue contain LPS and CLI-095 compound, the darker the color the higher the CLI-095 concentration and the bigger influence of the compound – cells are less inflamed and more similar to the positive control. Each well contains 49 microscopic images as presented in the middle of the figure. On the right, we present examples of final microscopic images and patches extracted from them. Top patches present cells from negative control and the bottom patches present cells from positive control.

drug discovery process where it accelerates screening of potential therapeutic compounds. Images are created using fluorescent reagents which mark specific cell structures, proteins, or DNA in a cell to measure its characteristic, e.g. chromatin in a nuclei [16], or perform more complicated tasks, like mode of action recognition [2].

Our work focuses on morphological changes in a microglia cell as a marker of inflammation. Activated microglia cells are key mediators of the chronic neuroinflammatory process, which is associated with the pathogenesis of many neurodegenerative disorders. During activation, apart from changes in the expression of surface antigens and production of pro-inflammatory factors, microglia also change their morphology, which is characterized by ameboid shape with many shortened processes [30,33]. Analyzing morphological changes can be, therefore, relevant in the determination of microglia activation and, as a consequence, can be useful for development of bioassays for the drug discovery process.

In our experiment, cells are treated with bacterial lipopolysaccharides (LPS) to induce inflammation and then treated with CLI-095 [17] to decrease it. To measure the influence of compounds, a biologist performs a quantitative analysis of cells to find how many cells are inflamed or active. Such analysis is performed semi-automatically because it requires a human to provide examples of both

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_23

active and inactive cells to a HCS analysis software[4]. Then, a linear machine learning model is trained to classify cells using morphological features, such as color intensity, cell shape and size. The main disadvantage of this procedure is an expensive and time-consuming step of manual cell labelling.

The goal of this work is to introduce a weakly-supervised method that can classify active and inactive cells, using a label of the entire image instead of single cells.

HCS images contain a variable number of cells which additionally react differently in response to the compound. Nevertheless, only one label, describing cell activity derived from a compound concentration, is assigned to an image. Problem with an input data point containing multiple instances, called the Multiple Instance Learning [10], often occurs in medical data, where it is too expensive and impractical to annotate details of an image. Similarly, in HCS images it is impossible to annotate every single cell for training purposes. In consequence, MIL is a perfect approach for such a problem, it is typically a weakly-supervised learning task with a goal of predicting a label for an entire image or a bag of data points. However, we want to broaden this idea by using MIL in a self-supervised setup to create an image representation that can be used to classify single instances. Our contributions are as follows:

- We propose a method for an instance level cell classification based on image level labels and apply it to HCS images.
- We introduce a novel position-aware Self-Supervised Multiple Instance Learning method (SSMIL) that combines Contrastive Learning and Multiple Instance Learning approaches.
- We demonstrate usability of our method with detailed tests conducted by both Machine Learning researchers and molecular biologists.

## 2 Related works

Firstly, we present current research in the field of High Content Screening, both using Deep Learning and classic approaches. Later, we summarize works on weakly-supervised and self-supervised learning with focus on Contrastive Learning.

### 2.1 High Content Screening

There are two main streams of High Content Screening image analysis: using classic hand-crafted features and deep learning methods. The first one utilizes morphological features that describe single cell and population-wide characteristics. Almost all manufacturers of HCS imagers provide image analysis software that uses such features, but noticeably more and more often Deep Learning modules are also provided. One of the most popular software is CellProfiler [5] which is free and open source. CellProfiler features were used for breast cancer detection

[4] https://www.perkinelmer.com/product/harmony-4-8-office-hh17000001

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_23

by creating hidden representations from an autoencoder model that reconstruct those hand-crafted features [21]. Also, they were used with regression models to predict phenotype [19, 35] and with Random Forests [31] for compound functional predictions. In recent years researchers also started to use Deep Learning methods for HCS analysis. [34] used a convolutional autoencoder to find abnormal cells by comparison to control images using deep learning features and [24] used variatonal autoencoders to show variation in cell phenotypes. At the same time, [14] used GANs for cell modelling. Convolutional neural networks are very widely used, [1, 3, 22, 23] and [29] even show that ImageNet-trained network can be successfully used to create features without fine-tuning. Additionally, some multi-scale CNN-based architectures were developed, [9, 13, 20]. Deep learning shows great results and the only disadvantage so far is lack of easy biological interpreatability of Deep Learning features.

### 2.2 Weakly-supervised learning

Weakly-supervised methods arise from the need to train models in case of lack of manual labels because, most often, only partial labels are available, e.g. in whole-slide images. CLAM [27] approaches this problem by performing instance-level clustering and then attention pooling in pathology classification. Similar problem to ours occurs in satellite images which are large and hard to annotate, it was tackled by use of stacked discriminative sparse autoencoder (SDSAE) [36]. Moreover, WELDON [11] automatically selects relevant regions by top-instance scoring.

### 2.3 Self-supervised learning

Self-supervised methods can create meaningful image representation without labels using so-called Contrastive Learning. Its main goal is to create similar representation for similar images, e.g. created by different augmenting of the same image. This is a base for Contrastive Predictive Coding (CPC) [28] which introduces contrastive loss and SimCLR [6, 7] which proposes NT-Xent loss. BYOL [15] transforms that idea by assigning online and target network, online network is trained continuously, while weights of the target network are the moving average of the first one. On the other hand, SwAV [4] utilizes swapped prediction mechanism and clustering, while SimSiam [8] simplifies contrastive idea by removing projection head from one of the paths and show that both representations (projected and not) are similar. So far, contrastive learning was used in MIL problems to create representation by training patch-level models [25, 26]. Our model utilizes it to compare image-level representation but pooling information from the model on the instance level to train a classifier for patches.

## 3 Data

HCS images are acquired from a high throughput microplate using an imager. Samples are prepared with multiple fluorescent dies which results in images with

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_23

various numbers of channels, even up to 6. In our experiment, cells were stained with two dyes: Hoechst 33342 and CellMask Deep Red Plasma Membrane. Microplate, or simply plate, has multiple wells organized into rows and columns. Each well consists of multiple images assigned to fields. In our case, the plate contains almost half a million cells with on average 150 cells per field and, in consequence, per image. Our plate is organized as follows (see Fig. 1): 2 columns contain controls (negative control does not contain any compound, positive control contains LPS which induces inflammation, called cell activity), 8 consecutive columns contain CLI-095 compound in decreasing concentration: from $1.45\mu M$ to $0.005\mu M$. CLI-095 is meant to decrease inflammation: the higher the concentration the smaller the inflammation (activity). In our training procedure, we treat negative control as having maximum concentration and positive control as having minimum concentration.

### 3.1 Data acquisition

BV2 microglia cells (Elabscience) were maintained in Dulbecco's modified Eagle's medium (DMEM) with 4.5 g/l glucose containing 5% (v/v) fetal bovine serum (Gibco), 1 mM sodium pyruvate (Gibco), 2mM L-glutamine (Gibco), 100 U/ml penicillin and 100 $\mu$g/ml streptomycin (Gibco) at 37◦ C in a humidified atmosphere containing 5% CO2. For study of morphological changes, BV2 cells were plated in assay medium (DMEM with 4.5 g/l glucose containing 1% (v/v) fetal bovine serum, 1 mM sodium pyruvate, 2mM L-glutamine , 100 U/ml penicillin and 100 $\mu$g/ml streptomycin) in poly-D-lysine coated CellCarrier Ultra 96-well plate (PerkinElmer) at density of 2000 cells/well. 24 hours later, cells were preincubated with different concentrations of CLI-095 in assay medium for 30 minutes, followed by stimulation with 0.1 $\mu$g/ml ultrapure LPS (Invivogen) and appropriate concentrations of CLI-095 in assay medium for further 24 hours. Final concentration of DMSO (CLI-095 solvent) was normalized to 0.0725%. After stimulation, cells were fixed with 4% formaldehyde for 25 minutes and washed with PBS 3 times. Then, nuclei were stained with Hoechst 33342 (TOCRIS) at the concentration of 5 $\mu$g/ml in PBS for 5 minutes and washed with PBS 3 times for 5 minutes. Then, cellular membrane was stained with CellMask Deep Red Plasma membrane (Invitrogen) at the dilution of 1:3000 in PBS for 10 minutes and washed with PBS 3 times for 5 minutes. Plates were then sealed and images were acquired with Operetta CLS (PerkinElmer) high content imaging system at conditions provided in Tab,1. Images were initially analyzed with use of Harmony software with Phenologic module (PerkinElmer).

### 3.2 Data labelling

Harmony is a software used along HCS Imager. It helps with image acquisition and experiment orchestration. Harmony has a basic image analysis function and can find and segment cells as well as provide morphological characteristics. Based on those, it uses Linear Regression (provided in the Phenologic module) to divide cells into up to 6 classes. Human operator selects training examples for each class

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_23

**Table 1.** Conditions of image acquisition.

| General settings | | |
|---|---|---|
| Autofocus | Two Peak | |
| Objective | 20x with water immersion, NA 1.0 | |
| Mode | Confocal | |
| Binning | 2 | |
| **Channel settings** | | |
| | Hoechst 3334 | CellMask Deep Red Plasma Membrane |
| Excitation | 355-385 nm | 615-645 nm |
| Emission | 430-500 nm | 655-760 nm |
| Time | 60 ms | 60 ms |
| Power | 40% | 20% |
| Height | $-6.0\mu m$ | $-11.0\mu m$ |

and software marks the rest of the data with those classes. We use Harmony labels as the ground truth. Cells might be in a transitional state which makes labels noisy and even a human operator is not able to distinguish between them. To assure correctness of the label, we take into account regression scores, which are also provided by Harmony, and use only cells with top and bottom 5% of values.

## 4 Methods

We introduce a novel position-aware method called SSMIL, which we compare to three baselines, two of which are supervised and one is self-supervised. Supervised methods include:

- Convolutional neural network (CNN) trained to classify cell activity based on image-level labels. We use ResNet-18 with single patches as an input and image labels.
- Attention-based Multiple Instance Learning Pooling (AbMILP) which aggregates all patches from an image to create one representation vector describing the entire image [18]. Additionally, we test its extension called Self-Attention Attention-based MIL Pooling (SA-AbMILP) [32].

A self-supervised baseline method is SimCLR [6] which is a contrastive method trained on patches containing single cells. The backbone of all methods is ResNet-18 to remove an influence of the CNN architecture.

### 4.1 Position-aware Self-Supervised Multiple Instance Learning method (SSMIL)

SSMIL, presented in Fig.2, combines Contrastive Learning and Multiple Instance Learning approaches. Following paragraphs describe each step of the model trained in an end-to-end manner.

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_23

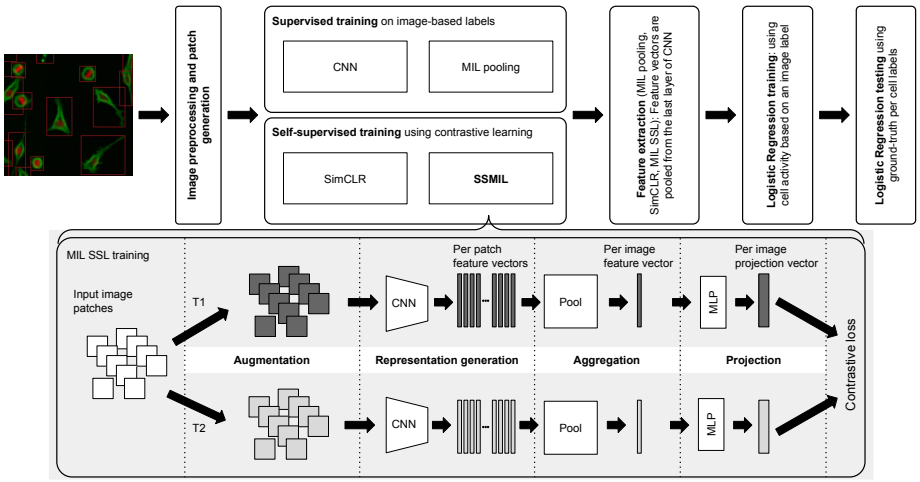


**Fig. 2.** Training pipeline: In the preprocessing step, we normalize images and remove those with quality issues, create patches and filter them due to size. In the training step, supervised models use compound's concentration as the output of a model and self-supervised models are trained in the contrastive manner. Then, we extract features from the ultimate (SSL) or penultimate (SL) layer of the CNN and we train Logistic regression using those features. LR is trained using cell activity based on the label of the entire image. Finally, we test our representation using ground-truth obtained with Harmony software. Additionally, we present the pipeline of the SSMIL model. Patches are augmented using two sets of transformations: $T1$ and $T2$, then passed through CNN to obtain patch representation which is aggregated using the pooling model to create representation of an entire image. Lastly this representation is passed through the projection module (MLP), and we calculate contrastive loss.

*Augmentation.* An image is split into patches generated for each previously detected cell. Then patches are transformed using two sets of augmentations, $T1$ and $T2$, to create two sets of patches.

*Representation generation.* Each patch is passed through a convolutional neural network to obtain its representation. In our setup, we use ResNet-18 without the classification layer. A patch representation is enriched by adding patch position at the end of the feature vector pooled from CNN. This information takes advantage of the fact that cells in the sample may influence each other, e.g. close cells might squeeze each other.

*Aggregation.* In the aggregation step, all patch feature vectors of the same image are pooled to create one feature vector representing an entire image. In this work, we use attention-based MIL pooling [18]:

$$z = \sum_{k=1}^{K} a_k h_k, \tag{1}$$

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_23

where $h_k$ is embedding of $kth$ patch and

$$a_k = \frac{\exp(\mathbf{w}^T \tanh(\mathbf{V}\mathbf{h}_k^T))}{\sum_{j=1}^{K} \exp(\mathbf{w}^T \tanh(\mathbf{V}\mathbf{h}_j^T))} \tag{2}$$

with $\mathbf{w}$ and $\mathbf{V}$ as trainable parameters.

*Projection.* Image feature vector is projected as according to [6] it is beneficial to use projection instead of feature vector. Projection corresponds to a 3-layer MLP.

*Contrastive loss.* Finally, we calculate contrastive loss using NT-Xent loss [6]:

$$L_{NT-Xent}(\mathbf{x}_i, \mathbf{x}_j) = -\log \frac{\exp(sim(\mathbf{x}_i, \mathbf{x}_j)/\tau)}{\sum_{k \neq i}^{2N} \mathbb{1}_{[k \neq i]} \exp(sim(\mathbf{x}_i, \mathbf{x}_k)/\tau} \tag{3}$$

where $\tau$ is temperature coefficient, $sim$ is cosine similarity, and $N$ is a batch size.

## 5 Experimental setup

Training and testing procedure for all models is presented in Fig.2 and is described in consecutive paragraphs.

*Train-test split.* We utilize the plate's structure to create a train-test split. The HCS plate is arranged in wells by rows and columns. Each well contains multiple fields (in our case 49). Our dataset has samples in 6 rows and we used 5 of them to select subsets. Rows are used in 5-fold split to create training (4 rows) and testing (1 row) datasets. Testing datasets are then filtered and we use only patches that have very high or low regression values, as given by Harmony software, to assure high confidence in labels used in testing. Our dataset is imbalanced because some cells die during the experiment due to induced inflammation.

*Data preprocessing and patch generation.* Cells are detected using Harmony software which also provides a bounding box enclosing a cell. We create $224 \times 224$ patches with the cell in the middle. Patches that are too close to the image border (112 pixels or less) and smaller than $500px^2$ are discarded to remove debris and cells that might not be correctly segmented. Finally, patches are resized to $112 \times 112$ pixels and normalized.

*Supervised and self-supervised training.* Supervised models are trained to classify patches (CNN) or images (AbMILP, SA-AbMILP) using activity labels based on the compound concentration. We assign active to columns with highest concentration of negative control and inactive to the lowest concentration columns and positive control. Self-supervised models (SimCLR, SSMIL) are trained in contrastive matter without any knowledge about the label. All models are trained

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_23

using Adam optimizer with learning rate $lr = 0001$. CNN is trained with batch size 128, SimCLR with batch size 1024. AbMILP, SA-AbMILP, and SSMIL are trained with batch size 8 which translates to on average 1200 patches. We use color jitter and rotation augmentations to minimize size changes of cells.

*Feature extraction.* Next, we create a patch representation from AbMILP, SA-AbMILP, SimCLR, and SSMIL models. The patch representation is given by feature vectors pooled from the last layer of CNN.

*Logistic Regression training.* In this step, we train Linear Regression using a patch representation to classify the cell activity. For training we use columns with the highest and lowest concentration to assure that the label is as correct as possible.

*Logistic Regression testing.* The last step of the procedure is testing. Linear regression models, trained in the previous step, are tested against labels obtained from Harmony+Phenologic software. CNN model is tested by using its direct prediction.

## 6 Results

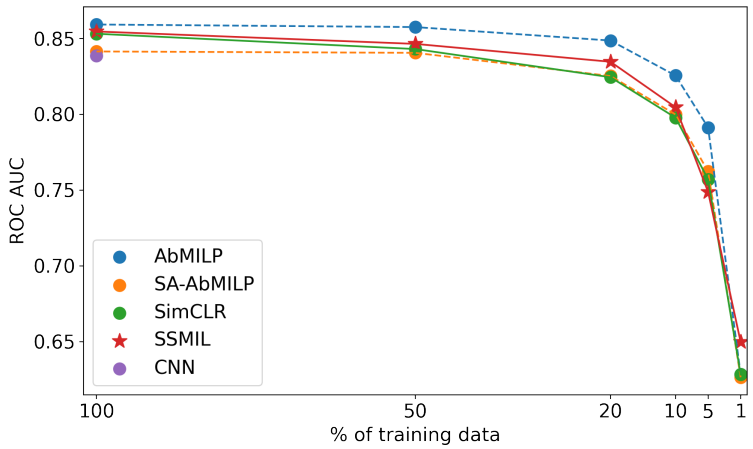


**Fig. 3.** Results: we present ROC AUC values for models trained with various fractions of training dataset. Our model SSMIL achieves on par results with supervised models and when trained on only 1% training data achieves 3.4% higher ROC AUC than other models, even supervised ones.

Results of our methods as well as baseline methods are presented in Fig. 3. We present ROC AUC (Receiver Operating Characteristic Area Under Curve)

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_23

values for Linear Regression models trained with 100%, 50%, 20%, 10%, 5% and 1% of the training data. Using 100% of training data SSMIL achieves ROC AUC of $0.8547 \pm 0.0129$ while AbMILP and SA-AbMILP have ROC AUC of $0.8593 \pm 0.0104$ and $0.8415 \pm 0.0147$, respectively. This shows that all models achieve very similar results and that by using self-supervised methods, we can obtain as good representation as in supervised setup. What is more, in all setups we achieve ROC AUC greater than 0.8 when we decrease size of training dataset to 20% of data or even 10% in case of SSMIL and supervised AbMILP. Worth noting is the fact that our model, SSMIL, achieves on average 3.4% better results than other models when trained only on 1% of data with SSMIL having ROC AUC of $0.6498 \pm 0.0102$ while other methods have ROC AUC of $0.6283 \pm 0.0076$, $0.6265 \pm 0.0125$, and $0.6283 \pm 0.0088$ (AbMILP, SA-AbMILP, and SimCLR respectively).

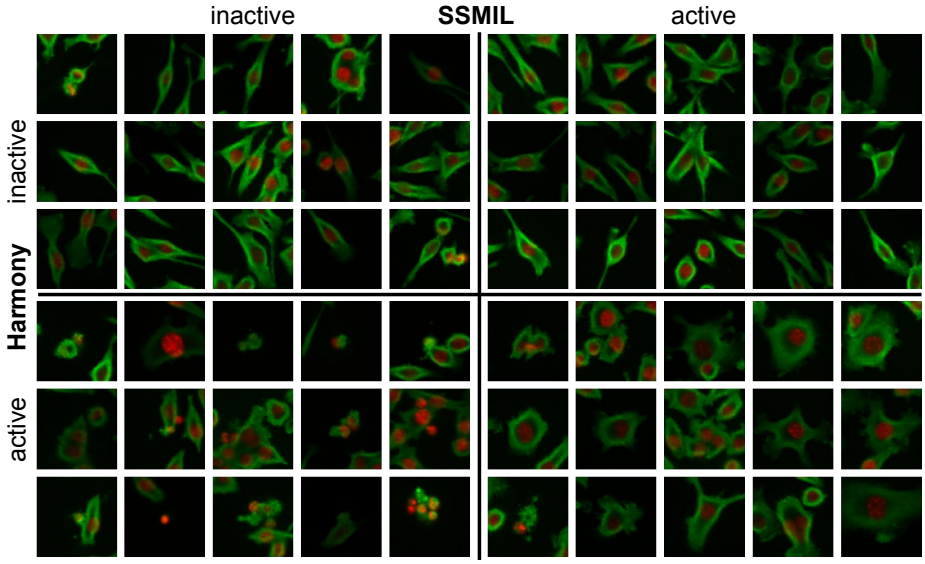


**Fig. 4.** Qualitative confusion matrix between Harmony labels and SSMIL predictions. SSMIL returns incorrect predictions for active cells mostly when those cells are destroyed or there are clumps of multiple cells. Similarly, the model incorrectly predicts inactive cells when they are too close to each other. In some cases we can see that cells labeled as inactive by Harmony are round and processes are not visible.

Fig.4 presents a qualitative confusion matrix which shows cells that were correctly and incorrectly predicted by the SSMIL model. We notice that SSMIL gives wrong predictions when faced with cells destroyed during the experiment. Additionally, the model can be confused by clumps of cells which can be mitigated by better segmentation which on its own can be a challenge when cells are touching.

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_23

## 7 Conclusions

We presented a weakly-supervised problem of cell classification using HCS data and compared 4 approaches to solve this problem in both supervised manner with CNN or MIL pooling, and using self-supervised methods, SimCLR and SSMIL proposed by us. We show that using either patch or image level Contrastive Learning provides representation as good as training a supervised model. Finally, we introduced a new contrastive learning method, SSMIL, that creates patch-level representations using aggregation and image-level labels and can be successfully trained even with an imbalanced data or a small number of training samples.

## References

1. Ando, D.M., McLean, C.Y., Berndl, M.: Improving phenotypic measurements in high-content imaging screens. bioRxiv (2017). https://doi.org/10.1101/161422, https://www.biorxiv.org/content/early/2017/07/10/161422
2. Bickle, M.: The beautiful cell: High-content screening in drug discovery. Analytical and bioanalytical chemistry **398**, 219–26 (09 2010). https://doi.org/10.1007/s00216-010-3788-3
3. Caicedo, J., McQuin, C., Goodman, A., Singh, S., Carpenter, A.: Weakly supervised learning of single-cell feature embeddings. vol. 2018, pp. 9309–9318 (06 2018). https://doi.org/10.1109/CVPR.2018.00970
4. Caron, M., Misra, I., Mairal, J., Goyal, P., Bojanowski, P., Joulin, A.: Unsupervised learning of visual features by contrasting cluster assignments. arXiv preprint arXiv:2006.09882 (2020)
5. Carpenter, A., Jones, T., Lamprecht, M., Clarke, C., Kang, I., Friman, O., Guertin, D., Chang, J., Lindquist, R., Moffat, J., Golland, P., Sabatini, D.: Cellprofiler: Image analysis software for identifying and quantifying cell phenotypes. Genome biology **7**, R100 (02 2006). https://doi.org/10.1186/gb-2006-7-10-r100
6. Chen, T., Kornblith, S., Norouzi, M., Hinton, G.: A simple framework for contrastive learning of visual representations (2020)
7. Chen, T., Kornblith, S., Swersky, K., Norouzi, M., Hinton, G.: Big self-supervised models are strong semi-supervised learners (2020)
8. Chen, X., He, K.: Exploring simple siamese representation learning. arXiv preprint arXiv:2011.10566 (2020)
9. Datta, K., Hossain, I., Choi, S., Saletore, V., Ambert, K., Godinez, W.J., Zhang, X.: Training multiscale-cnn for large microscopy image classification in one hour. High Performance Computing p. 463–477 (2019). https://doi.org/10.1007/978-3-030-34356-9_35, http://dx.doi.org/10.1007/978-3-030-34356-9_35
10. Dietterich, T.G., Lathrop, R.H., Lozano-Pérez, T.: Solving the multiple instance problem with axis-parallel rectangles. Artificial Intelligence **89**(1), 31–71 (1997). https://doi.org/https://doi.org/10.1016/S0004-3702(96)00034-3, https://www.sciencedirect.com/science/article/pii/S0004370296000343
11. Durand, T., Thome, N., Cord, M.: Weldon: Weakly supervised learning of deep convolutional neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (June 2016)

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_23

12. Giuliano, K., DeBiasio, R., Dunlay, R., Gough, A., Volosky, J., Zock, J., Pavlakis, G., Taylor, D.: High-content screening: A new approach to easing key bottlenecks in the drug discovery process. Journal of Biomolecular Screening - J BIOMOL SCREEN **2**, 249–259 (06 1997). https://doi.org/10.1177/108705719700200410
13. Godinez, W.J., Hossain, I., Lazic, S.E., Davies, J.W., Zhang, X.: A multi-scale convolutional neural network for phenotyping high-content cellular images. Bioinformatics **33**(13), 2010–2019 (02 2017). https://doi.org/10.1093/bioinformatics/btx069, https://doi.org/10.1093/bioinformatics/btx069
14. Goldsborough, P., Pawlowski, N., Caicedo, J.C., Singh, S., Carpenter, A.E.: Cytogan: Generative modeling of cell images. bioRxiv (2017). https://doi.org/10.1101/227645, https://www.biorxiv.org/content/early/2017/12/02/227645
15. Grill, J.B., Strub, F., Altché, F., Tallec, C., Richemond, P.H., Buchatskaya, E., Doersch, C., Pires, B.A., Guo, Z.D., Azar, M.G., et al.: Bootstrap your own latent: A new approach to self-supervised learning. arXiv preprint arXiv:2006.07733 (2020)
16. Haney, S., Lapan, P., Pan, J., Zhang, J.: High-content screening moves to the front of the line. Drug discovery today **11**, 889–94 (11 2006). https://doi.org/10.1016/j.drudis.2006.08.015
17. Ii, M., Matsunaga, N., Hazeki, K., Nakamura, K., Takashima, K., Seya, T., Hazeki, O., Kitazaki, T., Iizawa, Y.: A novel cyclohexene derivative, ethyl (6r)-6-[n-(2-chloro-4-fluorophenyl)sulfamoyl]cyclohex-1-ene-1-carboxylate (tak-242), selectively inhibits toll-like receptor 4-mediated cytokine production through suppression of intracellular signaling. Molecular Pharmacology **69**(4), 1288–1295 (2006). https://doi.org/10.1124/mol.105.019695, https://molpharm.aspetjournals.org/content/69/4/1288
18. Ilse, M., Tomczak, J., Welling, M.: Attention-based deep multiple instance learning. In: Dy, J., Krause, A. (eds.) Proceedings of the 35th International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 80, pp. 2127–2136. PMLR (10–15 Jul 2018), https://proceedings.mlr.press/v80/ilse18a.html
19. Janosch, A., Kaffka, C., Bickle, M.: Unbiased phenotype detection using negative controls. SLAS DISCOVERY: Advancing the Science of Drug Discovery **24**(3), 234–241 (2019). https://doi.org/10.1177/2472555218818053, https://doi.org/10.1177/2472555218818053, pMID: 30616488
20. Janssens, R., Zhang, X., Kauffmann, A., Weck, A., Durand, E.: Fully unsupervised deep mode of action learning for phenotyping high-content cellular images (07 2020). https://doi.org/10.1101/2020.07.22.215459
21. Kandaswamy, C., Silva, L.M., Alexandre, L.A., Santos, J.M.: High-content analysis of breast cancer using single-cell deep transfer learning. Journal of Biomolecular Screening **21**(3), 252–259 (2016). https://doi.org/10.1177/1087057115623451, https://doi.org/10.1177/1087057115623451, pMID: 26746583
22. Kensert, A., Harrison, P., Spjuth, O.: Transfer learning with deep convolutional neural networks for classifying cellular morphological changes. SLAS DISCOVERY: Advancing Life Sciences RD **24**, 247255521881875 (01 2019). https://doi.org/10.1177/2472555218818756
23. Kraus, O., Grys, B., Ba, J., Chong, Y., Frey, B., Boone, C., Andrews, B.: Automated analysis of high-content microscopy data with deep learning. Molecular Systems Biology **13**, 924 (04 2017). https://doi.org/10.15252/msb.20177551
24. Lafarge, M.W., Caicedo, J.C., Carpenter, A.E., Pluim, J.P., Singh, S., Veta, M.: Capturing single-cell phenotypic variation via unsupervised representation learning. In: Cardoso, M.J., Feragen, A., Glocker, B., Konukoglu, E.,

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_23

Oguz, I., Unal, G., Vercauteren, T. (eds.) Proceedings of The 2nd International Conference on Medical Imaging with Deep Learning. Proceedings of Machine Learning Research, vol. 102, pp. 315–325. PMLR (08–10 Jul 2019), https://proceedings.mlr.press/v102/lafarge19a.html
25. 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 (CVPR). pp. 14318–14328 (June 2021)
26. Lu, M.Y., Chen, R.J., Wang, J., Dillon, D., Mahmood, F.: Semi-supervised histology classification using deep multiple instance learning and contrastive predictive coding (2019)
27. Lu, M.Y., Williamson, D.F.K., Chen, T.Y., Chen, R.J., Barbieri, M., Mahmood, F.: Data efficient and weakly supervised computational pathology on whole slide images (2020)
28. van den Oord, A., Li, Y., Vinyals, O.: Representation learning with contrastive predictive coding (2019)
29. Pawlowski, N., Caicedo, J.C., Singh, S., Carpenter, A.E., Storkey, A.: Automating morphological profiling with generic deep convolutional networks. bioRxiv (2016). https://doi.org/10.1101/085118, https://www.biorxiv.org/content/early/2016/11/02/085118
30. Perry, V.H., Nicoll, J.A.R., Holmes, C.: Microglia in neurodegenerative disease. Nature Reviews Neurology **6**(4), 193–201 (Mar 2010). https://doi.org/10.1038/nrneurol.2010.17, https://doi.org/10.1038/nrneurol.2010.17
31. Rose, F., Basu, S., Rexhepaj, E., Chauchereau, A., Nery, E., Genovesio, A.: Compound functional prediction using multiple unrelated morphological profiling assays. SLAS TECHNOLOGY: Translating Life Sciences Innovation **23**, 247263031774083 (11 2017). https://doi.org/10.1177/2472630317740831
32. Rymarczyk, D., Borowa, A., Tabor, J., Zieliński, B.: Kernel self-attention in deep multiple instance learning (2021)
33. Sarkar, S., Malovic, E., Sarda, D., Lawana, V., Rokad, D., Jin, H., Anantharam, V., Kanthasamy, A., Kanthasamy, A.G.: Characterization and comparative analysis of a new mouse microglial cell model for studying neuroinflammatory mechanisms during neurotoxic insults. NeuroToxicology **67**, 129–140 (Jul 2018). https://doi.org/10.1016/j.neuro.2018.05.002, https://doi.org/10.1016/j.neuro.2018.05.002
34. Sommer, C., Hoefler, R., Samwer, M., Gerlich, D.: A deep learning and novelty detection framework for rapid phenotyping in high-content screening. bioRxiv (09 2017). https://doi.org/10.1101/134627
35. Way, G., Kost-Alimova, M., Shibue, T., Harrington, W., Gill, S., Piccioni, F., Becker, T., Hahn, W., Carpenter, A., Vazquez, F., Singh, S.: Predicting cell health phenotypes using image-based morphology profiling (07 2020). https://doi.org/10.1101/2020.07.08.193938
36. Yao, X., Han, J., Cheng, G., Qian, X., Guo, L.: Semantic annotation of high-resolution satellite images via weakly supervised learning. IEEE Transactions on Geoscience and Remote Sensing **54**(6), 3660–3671 (2016). https://doi.org/10.1109/TGRS.2016.2523563

ICCS Camera Ready Version 2022
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-08751-6_23