

# Unsupervised Deep Learning of Bright Field Images for Apoptotic Cell Classification

**Zhuo Zheng**

South China Normal University

**Beini Sun**

South China Normal University

**Siqi He**

South China Normal University

**Guanchen Wang**

South China Normal University

**Chenyang Bi**

South China Normal University

**Tongsheng Chen** (✉ [chentsh@scnu.edu.cn](mailto:chentsh@scnu.edu.cn))

South China Normal University

---

## Research Article

**Keywords:** Image classification, Deep learning, Apoptosis, Autoencoder

**Posted Date:** August 19th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1956878/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**Additional Declarations:** No competing interests reported.

---

**Version of Record:** A version of this preprint was published at Signal, Image and Video Processing on April 24th, 2023. See the published version at <https://doi.org/10.1007/s11760-023-02592-1>.

# Abstract

The classification of apoptotic and living cells is significant in drug screening and treating various diseases. Conventional supervised methods require a large amount of prelabelled data, which is often costly and consumes immense human resources in the biological field. In this study, unsupervised deep-learning algorithms were used to extract cell characteristics and classify cells. A model integrating a convolutional neural network and an auto-encoder network was utilised to extract cell characteristics, and a hybrid clustering approach was employed to obtain cell feature clustering results. Experiments on both public and private datasets revealed that the proposed unsupervised strategy performs well in cell categorisation. For instance, in the public dataset, our method obtained a precision of 96.72% on only 1000 unlabelled cells. To the best of our knowledge, this is the first time unsupervised deep learning has been applied to distinguish apoptosis and live cells with high accuracy.

## 1 Introduction

Apoptosis is the cell-independent orderly death that occurs to resist external stimulation and maintain the homeostasis of the internal environment, which is often referred to as programmed death. Unlike other methods of cell death, apoptosis is a self-protection mechanism, activated, expressed, and regulated by specific genes [1, 2]. The term apoptosis was first coined by Kerr et al. in 1972 to describe a new morphological feature of cell death [3]. *Caenorhabditis elegans*' programmed cell death has led to a better understanding of apoptosis in mammals [4]. Physiologically, apoptosis plays a crucial role in the growth, development, and evolution of organisms [5]. The process of apoptosis maintains homeostasis and the dynamic balance of cell number in the body, but it can also be used as a defence mechanism to eliminate unnecessary or unwanted cells [6]. Appearing in cell apoptosis, cell shrinkage, smaller volume, nuclear enrichment, nuclear membrane nucleoli, DNA fragmentation, then the cell cleaves into apoptotic corpuscles, which are formed by the cell membrane-enclosed cytoplasm, organelles and broken nucleus, and eventually, the apoptotic body are recognized around the macrophage, which is swallowed, resulting in degradation. The cell membrane structure is complete during apoptosis, no contents are spilled, no cytokines are released, the duration is short, and there is no inflammatory reaction around. There are several signalling pathways involved in apoptosis. Apoptosis in cells can be triggered by either the caspase-mediated extrinsic or intrinsic pathways. Both pathways converge to activate effector apoptotic caspases, ultimately resulting in morphological and biochemical cellular alterations and characteristics of apoptosis [7].

Apoptosis is a rational and active decision made to sacrifice specific cells for the greater benefit of an organism [8]. Detection of apoptotic cell death in cells and tissues has become of paramount importance in many fields of modern biology, including studies of embryonic development, degenerative disease, and cancer biology [9]. TEM is regarded as the gold standard for conforming apoptosis. This is because the categorisation of an apoptotic cell is irrefutable if the cell exhibits certain ultrastructural morphological characteristics [10]. Cost, time, and only being able to analyze a small area at a time are the major disadvantages of TEM. The TUNEL (Terminal dUTP Nick End-Labeling) method is used to assay the

endonuclease cleavage products by enzymatically end-labelling the DNA strand breaks [11]. Terminal transferase is used to add labelled UTP to the 3'-end of the DNA fragments. The dUTP can then be labelled with a variety of probes to allow detection by light microscopy, fluorescence microscopy or flow cytometry. These techniques are costly, potentially phototoxic, and may even interfere with the cell death process itself [12, 13]. As a result, novel alternative approaches to studying apoptosis and distinguishing it from living cells are required.

In 1998, Lecun proposed the convolutional neural network (CNN) and demonstrated excellent performance in image recognition [14]. Following that, some improved CNNs with deeper layers were designed and achieved satisfactory image recognition results, including AlexNet, VGGNet, GoogLeNet, and ResNet [15, 16]. CNNs mainly include the input layer, output layer, convolutional layer and pooling layer. The most important layers in CNNs are the convolutional layers themselves. Different convolution kernels are filtered without interfering with each other, from an input image or a feature map output from the previous layer, each filter effectively extracts a particular type of feature. In recent years, CNNs have been widely applied to medical imaging analysis [17, 18]. Machine learning algorithms have been demonstrated to be successful in classifying apoptotic and live cells [19]. To achieve successful performance in supervised learning, such as with CNNs, a large amount of high-quality human-labelled data from the target domain is required. In the bioimage informatics domain, structured and machine-readable labelling is still uncommon and generally very expensive to produce. To address the issue of collecting high-quality labelled data while utilizing a large amount of existing unlabelled data, we applied an unsupervised deep learning method to our model.

As an unsupervised pre-training method for artificial neural networks (ANN), the autoencoder was originally called the "auto-associative learning module" in the 1980s [20]. As a method of learning features without supervision, autoencoders have been widely adopted. In order to improve their performance, autoencoders' outputs are often used as inputs into other networks and algorithms. There are generally two modules in an autoencoder: an encoder and a decoder. The encoding module encodes the input signals into a latent space, while the decoder module transforms signals from the latent space back into the original domain. The main applications of autoencoders are dimensionality reduction and feature extraction [21–23]. In this study, we used a deep learning unsupervised classification model that integrated a CNN and an auto-encoding network known as the constraint convolutional auto-encoder (CCAIE) to extract cell characteristics, and a hybrid clustering approach was applied to obtain the final clusters [24]. Certain changes were made to improve the performance of CCAIE, such as adding a new loss function to aid model convergence and improve the network fundamental structure. Experiments on public and private datasets revealed that this unsupervised method performed well in cell categorisation. We also created a bright field imaging dataset of HeLa cells and calculated the accuracy of classifying apoptotic cells to confirm the effectiveness of our method. To verify the generalisation of the network, we tested it on a public dataset with the same parameters. This method yielded positive results for both datasets.

The remainder of this paper is organized as follows. In Section 2, we introduce the main method, including the data pre-processing, the CCAE neural network model structure, and the hybrid clustering strategy respectively. In Section 3, we exhibit the result of the main method used on datasets. Furthermore, details such as the experimental environment's configuration, the specific configuration of the experimental model, and the performance evaluation of each model are provided. Conclusions and discussions are discussed in Section 4.

## 2 Dataset And Main Method

### 2.1 Cells and materials

Human cervical cancer cells (HeLa) were cultured in DMEM (Gibco, USA) supplemented with 10% foetal bovine serum (Sijiqing, China) in a 37 °C incubator containing 5% CO<sub>2</sub>. Staurosporine (STS) was purchased from Sigma-Aldrich (United). The cells were incubated with 1 µM STS for 1, 2, 4, and 8 h, respectively, to induce apoptosis and subsequently imaged bright fields using a Zeiss (Axio Observer 7, Germany). Using fluorescence microscopy to demonstrate that apoptosis was successfully induced.

### 2.2 Image pre-processing

Generally, bright-field images of cells are susceptible to interference, such as noise, uneven grayscale, and uneven image contrast. Cell image pre-processing comprises global equalisation transformation, threshold processing, adaptive binarization, and denoising to obtain high-quality images. Gaussian blur filter and contrast equalisation are initially used to remove high-frequency noise and boost image contrast. Subsequently, adaptive binarization processing is used to maximise differences in the representation of foreground and background and improve image quality. Finally, we can emphasise the morphological characteristics of the cells by minimising the background noise in the images, as shown in Fig. 1.

### 2.3 Deep learning model

This section introduces the CCAE convolutional auto-encoder structure and the modifications we made to improve the performance of the model. CCAE is a variant of the autoencoder (AE) model. The AE coding block is replaced by convolutional layers in CCAE, allowing CCAE to effectively compress images, and a constraint is applied to the hidden parameter to assist the model in learning more effective features.

The AE is an efficient method for extracting valid information from raw data [25]. The main information is preserved while the input dimension is reduced by encoding and decoding blocks composed of multiple fully connected layers. Normally, the encoding block can be regarded as a

nonlinear mapping that maps the input data  $x$  to the latent vector parameter  $y$ . For instance,

$$y = f_{\text{encode}}(x) = \sigma(Wx + b)$$

where  $\sigma(x)$  denotes the nonlinear activation function, and  $W$  and  $b$  are the weight and bias of the encoding block, respectively. The decoding block, which is another nonlinear mapping, maps the latent vector parameter  $y$  back to the reconstructed data  $x'$ . By adjusting the weights in the encoding and decoding blocks, the AE minimises the difference between the reconstructed data  $x'$  and input data  $x$ .

The convolutional operation is an effective method of extracting image features. For instance, AlexNet [26] and Lenet-5 networks [27], as variations of the CNNs, enhanced the image classification efficiency significantly. It has been shown that applying a constraint on the hidden parameter  $y$  helps the AE to learn more effective features [28]. Therefore, in VAE, the hidden parameters are usually sampled using a Gaussian distribution. In addition, Aytekin et al. [29] sampled the latent vector parameters from the unit ball space.

To assist the model convergence and enhance the network fundamental structure, we used triplet loss instead of L2 loss to optimise the effective features of cell images. The triplet loss pulls in the positive sample pair and pushes out the negative sample pair such that the same label image features can be aggregated in the feature space [30]. In this algorithm, the anchor picture is treated as a positive image, whereas all other photos are treated as negative images. Extracting all morphological features of cells is redundant for cell classification and will lead to a decrease in classification accuracy. In other words, the model must retain valid features while discarding invalid ones. In this model, the latent vector parameter  $y$  was sampled from a Gaussian distribution with a mean value of 0. The constraint is transformed into an additional term in the loss function,

$$\text{loss} = 1/M \sum_i^M \left\{ \|x_i - x'_i\|_2^2 + \sum_i^N [-\|f(x_i^a) - f(x_i^n)\|_2^2 + \alpha]_+ + 1/L \sum_j^L y_{ij}^2 \right\}$$

2

where  $M$  denotes the number of cell images,  $N$  means the number of triplets,  $L$  is the length of the latent vector,  $x_i$  is the input image,  $x'_i$  is the decoding data,  $x_i^a$  is the anchor image,  $x_i^n$  means the negative image and  $\alpha$  is the margin between positive and negative pairs. The structure of the model and the parameter setting is shown in Fig. 2.

## 2.4 Hybrid clustering algorithm

This section introduces a new clustering strategy known as the hybrid clustering algorithm for cell clustering. The single clustering method groups subsamples from a single perspective, which leads to misclassification. The sub-samples can be grouped from a comprehensive perspective, which is multi-perspective by considering the clustering results of various types of models; thus, the clustering results are more reliable [31]. The encoded data were used to divide the cell data into two categories (apoptotic and live cells) using different clustering methods with different clustering effects. The hybrid clustering algorithm, a bagging algorithm based on the multi-clustering model, is then used to secure the optimum

clustering results. In other words, a cell was considered apoptotic if it was considered apoptotic by all of the clustering techniques I used.

Three traditional clustering methods were used in this study: the Euclidean distance k-means [32], agglomerative

(AGG) [33], and balanced iterative reducing and clustering using hierarchies (BIRCH) [34] algorithms. The k-means algorithm divides dataset samples into k cluster classes using different distance formulas. The cluster centre was obtained by initialising the mean vector at the beginning. The greedy strategy minimises the distance between the sample and the cluster centre, and the cluster centre is updated. Finally, clustering results were obtained. The AGG and BIRCH algorithms belong to the hierarchical clustering algorithm, which divides the dataset at different levels into a tree-like structure. The AGG algorithm is a bottom-up aggregation strategy. First, each sample in the dataset is considered an initial cluster; then, at each step of the algorithm, the two clustering clusters that are closest to each other are found to merge. The merging process was repeated until the pre-set number of cluster clusters was reached. The BIRCH algorithm uses a clustering feature (CF) tree to perform hierarchical clustering. The algorithm constructs a CF tree based on input data. Then, the clustering algorithm and outlier processing on the leaf nodes were performed. At the end of the clustering process, each leaf node becomes a cluster of sample sets. Although the multi-clustering model-based bagging algorithm generates more rejection data, the hybrid clustering algorithm can significantly improve clustering accuracy, as described in the following sections. Figure 3 provides a brief overview of the main methods used for cell classification.

### 3. Results

In this section, we provide evaluation indicators for several models. The transfer of the evaluation method with supervised learning to the unsupervised task is the highlight here. All our models used an Adam [35] optimiser with a learning rate of  $3 \times 10^{-4}$  training 1000 epochs on datasets. The experimental system environment used for learning and testing was as follows: The operating system is Linux Ubuntu 18.04.5; The hardware consists of an Intel Xeon CPU E5-2690v3 2.60GHz, 64 GB memory, two NVIDIA GeForce GTX 3090 GPU; Tensorflow 2.4.0 deep learning framework. A table is provided to demonstrate the performance of several datasets to show the difference between the existing supervised methods and our model.

Table 1 reveals that CCAE achieves higher accuracy than the supervised method without any pre-labelled training dataset. In addition, the model can accurately determine whether the cells are apoptotic when the morphology of the cells changes slightly.

#### 3.1 Effectiveness of the hybrid clustering algorithm

In this subsection, we compare the results of the single clustering model with the hybrid clustering model, as summarized in Table 2, to demonstrate the usefulness of the hybrid clustering technique. Table 2

reveals that the accuracy of the hybrid clustering model is at least 1.08% higher than that of the single clustering model at the cost of at most 6.40% rejected subsamples.

Table 1

Cell image classification results on our dataset and public dataset. The classification results of current supervised deep learning method for apoptotic and live cells was displayed in the first row of the table. The classification results of the CCAE on 1000 randomly chosen images from the same dataset were displayed in the second row. The remaining rows show the classification results of CCAE on our own dataset. Our method yielded positive results for all datasets. R.r is the Reject rate, which is the percentage of the total number of rejections after the hybrid clustering mechanism.

Method	Training sets(size)	Testing sets(size)	Alive prediction (%)	Apoptosis prediction (%)	R.r (%)
Verduijn et al.	DLAN (35000)	DLAN (70000)	87.90	81.80	×
The CCAE	None	DLAN (1000)	91.00	96.72	6.40
The CCAE	None	Hela treated with STS for 8h (1000)	97.32	98.42	2.74
The CCAE	None	Hela treated with STS for 4h (1000)	98.45	98.20	0.80
The CCAE	None	Hela treated with STS for 2h (1000)	97.05	97.98	2.20
The CCAE	None	Hela treated with STS for 1h (1000)	91.95	92.07	2.14

Table 2

Comparison of independent testing performance between single clustering model and hybrid clustering model. On the public dataset and our dataset, we tried the k-means, BIRCH, and AGG and hybrid clustering methods individually. Experimental results on the both datasets employed show that, compared to single clustering methods, hybrid clustering method generally achieves the best classification accuracy. By considering the clustering result of various kinds of models, one can group subsamples from a comprehensive view, that is a multi-perspective of view, and thus the clustering result is more reliable.

Method	Testing sets(size)	Alive prediction (%)	Apoptosis prediction (%)	R.r(%)
The $CCAE_{k-means}$	DLAN(1000)	89.79	94.69	×
The $CCAE_{BIRCH}$	DLAN(1000)	90.12	92.53	×
The $CCAE_{AGG}$	DLAN (1000)	88.42	89.17	×
The $CCAE$	DLAN (1000)	91.00	96.72	6.40
The $CCAE_{k-means}$	HeLa treated with STS for 8h (1000)	95.33	98.25	×
The $CCAE_{BIRCH}$	HeLa treated with STS for 8h (1000)	92.19	90.73	×
The $CCAE_{AGG}$	HeLa treated with STS for 8h (1000)	84.43	94.96	×
The $CCAE$	HeLa treated with STS for 8h (1000)	97.32	98.42	2.74

## 3.2 Analysis of CCAE

This section delves deeper into the efficacy of CCAE. The triplet loss-function restriction is discussed in detail. Our triplets consist of two matching cell bright field images and a non-matching cell bright field image, and the loss aims to separate the positive pair from the negative pair by a distance margin. Particularly, we seek an embedding  $f(x)$ , from an image  $x$  into a feature space such that the squared distance between all cells of the same identity, independent of imaging conditions, is small, whereas the squared distance between a pair of cell images from different images is large. In this algorithm, the anchor picture is treated as a positive image, whereas all other photos within the margin are treated as negative images. We compare triplet loss to l2 loss as shown in Table 3.

The strong prior operation is favourable for model learning features. We used the t-SNE (t-Distributed Stochastic Neighbour Embedding) [36] visualization method to directly show the improvement brought by strong prior operation. The technique is a variation of Stochastic Neighbour Embedding [37]. T-SNE can

capture much of the local structure of high-dimensional data while also revealing global structures such as the presence of clusters at multiple scales. The results of the t-SNE visualization are shown in Fig. 4.

A constraint is added to the encoding network so that the generated latent variables roughly follow the standard normal distribution, potentially improving recognition accuracy [38]. A constraint on the latent vector is a mechanism for "forgetting" cell characteristics. We add Gaussian constraint to the encoding data to help our model learn more effective features.

## 4. Conclusions And Discussions

In this study, we combine convolutional neural networks and an autoencoder to extract cell characteristics and use the hybrid clustering algorithm to cluster cells. The encoder utilises convolutional layers to learn cell features from images and the decoder uses deconvolutional layers to reconstruct cell images. Based on FaceNet [30], which uses triplet loss to learn highly discriminating face features, we introduce triplet loss to guide the model to learn recognizable cell features. We employed a generative model instead of a discriminative model to extract cell structural information in its entirety. The goal of the discriminant model is to locate a high-dimensional representation of data in a high-dimensional space and thereafter transfer them to a low-dimensional space for classification and discrimination. This may contain data with missing structural properties. In contrast, the generation model can compress the data dimension while the generation module maintains the structural integrity of the data. This is a more sensible method of data compression with greater resilience and interpretability for downstream data matching and classification.

Table 3

Comparison of CCAE performance with L2 loss and triplet loss. In the Table, the first and the second row report the classification results on DLAN dataset by using CCAE with different loss function. Classification results of CCAE with different loss on our dataset are present in the remaining rows. On the same dataset, the triplet loss function not only has greater recognition accuracy than l2 loss, but also rejects less photographs, as seen in the table.

Method	Testing sets(size)	Alive prediction (%)	Apoptosis prediction (%)	R.r(%)
The <i>CCA</i> E with L2 loss	DLAN (1000)	88.17	92.10	12.50
The <i>CCA</i> E with triplet loss	DLAN (1000)	91.00	96.72	6.40
The <i>CCA</i> E with L2 loss	HeLa treated with STS for 8h (1000)	91.36	93.42	14.73
The <i>CCA</i> E with triplet loss	HeLa treated with STS for 8h (1000)	97.32	98.42	2.74

Prior study has shown that supervised deep learning performs well in discriminating cell images. Verduijn's method used a training set (~ 35,000 cells) to tune the model towards its new objective of detecting and discriminating between alive, apoptotic and necroptotic cancer cells [19]. It is possible to use deep learning algorithms such as the convolutional neural network for image classification, however, these methods require an extensive amount of training data that is well-annotated. Compared with the existing methods, our algorithm only uses 1000 randomly selected cell images from Verduijn's dataset and achieves higher accuracy. Although our method can only distinguish living and apoptotic cells, it does not require any annotated images compared with other deep learning models and can achieve a considerably high classification accuracy.

This is the first study to look into the potential of unsupervised deep learning methods for apoptotic cell classification. The triplet loss function is introduced into the CCAE loss module, which uses image similarity to enhance feature extraction ability. We demonstrate the performance by the published dataset, and the result shows that our proposed method can achieve a high classification accuracy rate. Compared with some existing supervised methods, our method is more applicable to domains in which it is difficult to obtain high-quality, balanced datasets. The approach described here is expected to have a significant impact on research in cancer diagnosis, drug screening, and cell characterization in general.

This study has several limitations that should be acknowledged. Our approach has only been tested on brightfield images of HeLa cells and digital holographic microscopy images of L929sAhFas cells so far. Furthermore, some cell deaths, such as ageing, necrosis, and ferroptosis, could help in a variety of applications, such as developing and screening new drugs and detecting dangerous toxins. It remains to be seen whether unsupervised deep learning approaches work well in other cell lines and cell death. We intend to investigate the effect of our method on a larger cell dataset in the future and make it applicable to multiple cell classification problems.

For future work, we will consider different network architectures, such as residual [40] and the inception [41] networks, which have been shown highly effective in finding latent structures and features. Furthermore, we plan to enlarge the capabilities and generality of the method and migrate the model to other domains.

## Declarations

**Ethical Approval** Not applicable.

**Competing interests** The authors declare no competing financial interest.

**Authors' contributions** Z-Z designed and performed experiments, analyzed data, and wrote the paper. BN-S provided cell pictures. SQ-H, GC-W and CY-B find relevant information and provide ideas. TS-C designed the study and planned the experiments.

**Funding** This work was supported by grants from the Key-Area Research and Development Program of Guangdong Province (Grant numbers: 2022B0303040003) and the National Natural Science Foundation of China (NSFC) (Grant numbers: 62135003 and 61875056).

**Availability of data and materials** The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

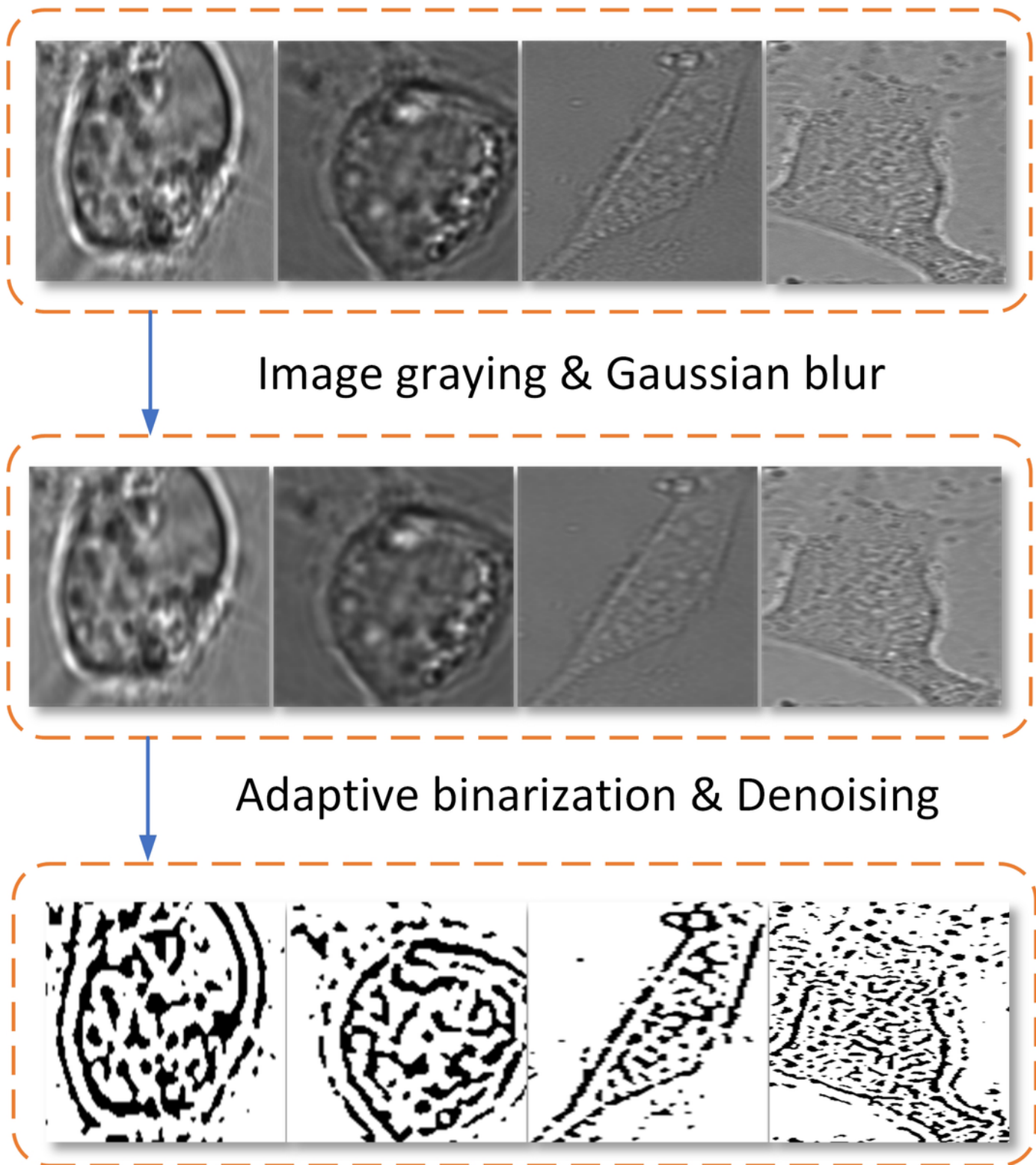
## References

1. Kaczanowski, S.: Apoptosis: its origin, history, maintenance and the medical implications for cancer and aging. *Phys. Biol.* **13**(3), 31001 (2016)
2. Ntuli, T.M.: Apoptosis and medicine. IntechOpen, New York (2012)
3. Kerr, J.F., Wyllie, A.H., Currie, A.R.: Apoptosis: a basic biological phenomenon with wideranging implications in tissue kinetics. *Brit. J. Cancer* **26**(4), 239–257 (1972).
4. Horvitz, H.R.: Genetic control of programmed cell death in the nematode *Caenorhabditis elegans*. In: Mihich, E., Schimke, R.T. (eds.) *Apoptosis*, pp. 1–13. Springer, Boston (1994)
5. Fuchs, Y., Steller, H.: Live to die another way: modes of programmed cell death and the signals emanating from dying cells. *Nat. Rev. Mol. Cell Bio.* **16**(6), 329–344 (2015)
6. Norbury, C.J., Hickson, I.D.: Cellular responses to DNA damage. *Annu Rev Pharmacol* **41**(1), 367–401 (2001)
7. Wong, R.S.: Apoptosis in cancer: from pathogenesis to treatment. *J. Exp. Clin. Canc. Res.* **30**(1), 1–14 (2011)
8. Kerr, J.F.: History of the events leading to the formulation of the apoptosis concept. *Toxicology* **181**, 471–474 (2002)
9. Willingham, M.C.: Cytochemical methods for the detection of apoptosis. *J. Histochem. Cytochem.* **47**(9), 1101–1109 (1999)
10. Elmore, S.: Apoptosis: a review of programmed cell death. *Toxicol. Pathol.* **35**(4), 495–516 (2007)
11. Ito, Y., Otsuki, Y.: Localization of apoptotic cells in the human epidermis by an in situ DNA nick end-labeling method using confocal reflectant laser microscopy. *J. Histochem. Cytochem.* **46**(6), 783–786 (1998)
12. Purschke, M., Rubio, N., Held, K.D., Redmond, R.W.: Phototoxicity of Hoechst 33342 in time-lapse fluorescence microscopy. *Photoch. Photobio. Sci.* **9**(12), 1634–1639 (2010)
13. Laissue, P.P., Alghamdi, R.A., Tomancak, P., Reynaud, E.G., Shroff, H.: Assessing phototoxicity in live fluorescence imaging. *Nat. Methods* **14**(7), 657–661 (2017)
14. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. *Proc. IEEE* **86**(11), 2278–2324 (1998)
15. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf Process Syst* **25**, 1097–1105 (2012)

16. Simonyan, K., Zisserman, A.; Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
17. Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G., Thrun, S., Dean, J.: A guide to deep learning in healthcare. *Nat. Med.* **25**(1), 24–29 (2019)
18. Kermany, D.S., Goldbaum, M., Cai, W., Valentim, C.C., Liang, H., Baxter, S.L., McKeown, A., Yang, G., Wu, X., Yan, F.: Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* **172**(5), 1122–1131 (2018)
19. Verduijn, J., Van der Meeren, L., Krysko, D.V., Skirtach, A.G.: Deep learning with digital holographic microscopy discriminates apoptosis and necroptosis. *Cell death discovery* **7**(1), 1–10 (2021)
20. Ballard, D.H.: Modular learning in neural networks. Paper presented at the 6th AAAI, Seattle, WA, 279–284 July 1987 (1987)
21. Varamin, A.A., Abbasnejad, E., Shi, Q., Ranasinghe, D.C., Rezaatofghi, H.: Deep auto-set: A deep auto-encoder-set network for activity recognition using wearables. Paper presented at proceedings of the 15th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, New York, USA, 246–253 November 2018 (2018)
22. Jia, G., Lam, H., Liao, J., Wang, R.: Classification of electromyographic hand gesture signals using machine learning techniques. *Neurocomputing* **401**, 236–248 (2020)
23. Rubio-Solis, A., Panoutsos, G., Beltran-Perez, C., Martinez-Hernandez, U.: A multilayer interval type-2 fuzzy extreme learning machine for the recognition of walking activities and gait events using wearable sensors. *Neurocomputing* **389**, 42–55 (2020)
24. Hou, Y.J., Xie, Z.X., Zhou, C.C.: An unsupervised deep-learning method for fingerprint classification: the ccae network and the hybrid clustering strategy. arXiv preprint arXiv:2109.05526 (2021)
25. Baldi, P.: Autoencoders, unsupervised learning, and deep architectures. Paper presented at proceedings of the 2011 International Conference on Unsupervised and Transfer Learning workshop, Washington, USA, 37–50 July 2011 (2011)
26. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf Process Syst* **25**, 1097–1105 (2012)
27. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. *Proc. IEEE* **86**(11), 2278–2324 (1998)
28. Kingma, D. P., Welling, M.: Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114 (2013)
29. Aytekin, C., Ni, X., Cricri, F., Aksu, E.: Clustering and unsupervised anomaly detection with  $l_2$  normalized deep auto-encoder representations. Paper presented at 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, 1–6 July 2018 (2018)
30. Schroff, F., Kalenichenko, D., Philbin, J.: Facenet: A unified embedding for face recognition and clustering. Paper presented at proceedings of the IEEE conference on computer vision and pattern recognition, Boston, USA, 815–823 June 2015 (2015)

31. Zhou, C., Gu, Y., Fang, G., Lin, Z.: Automatic Morphological Classification of Galaxies: Convolutional Autoencoder and Bagging-based Multiclustering Model. *Astron. J.* **163**(2), 86 (2022)
32. MacQueen, J.: Some methods for classification and analysis of multivariate observations. Paper presented at proceedings of the 5th Berkeley symposium on mathematical statistics and probability, Oakland, USA, 281–297 July 1966 (1967)
33. Murtagh, F.: A survey of recent advances in hierarchical clustering algorithms. *Comput J* **26**(4), 354–359 (1983)
34. Zhang, T., Ramakrishnan, R., Livny, M.: BIRCH: an efficient data clustering method for very large databases. *SIGMOD Rec* **25**(2), 103–114 (1996)
35. Kingma, D.P., Adam, B.J.: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
36. Van der Maaten, L., Hinton, G.: Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**(11), 2579–2605 (2008)
37. Hinton, G.E.: Training products of experts by minimizing contrastive divergence. *Neural Comput.* **14**(8), 1771–1800 (2002)
38. Kingma, D.P., Welling, M.: Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114 (2013)
39. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. Paper presented at 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, USA, 770–778 June 2016 (2016)
40. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. Paper presented at 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, USA, 1–9 June 2015 (2015)

## Figures



**Figure 1**

Preprocessing of HeLa cells treated with STS for 8 h and control. An overview of changes were made to bright filed images of apoptotic and living cells. Image graying and Gaussian blur filter can remove high-frequency noise and boost image contrast. Adaptive binarization processing and denoising is used to maximize differences in the representation of foreground and background and improve image quality

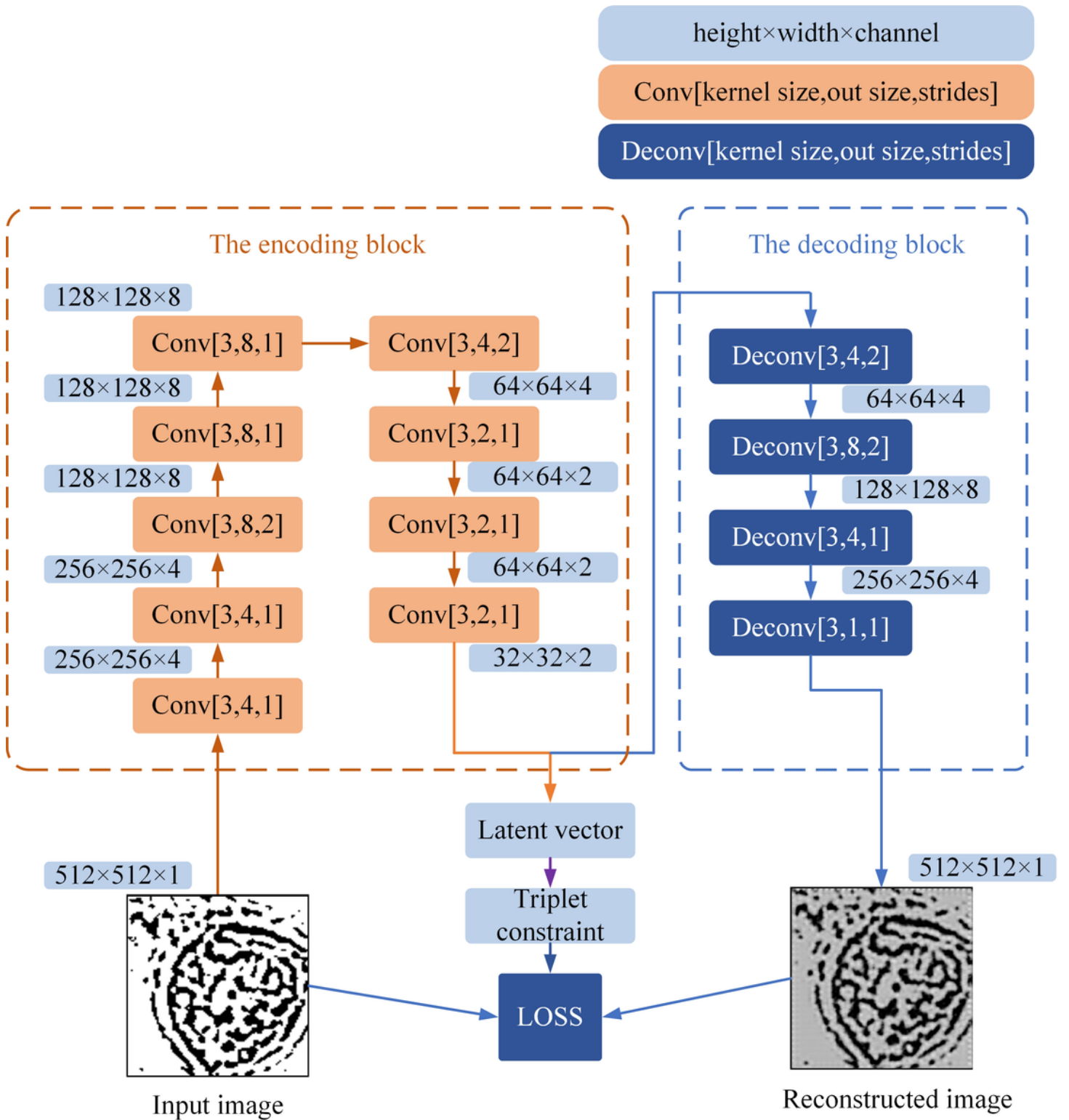
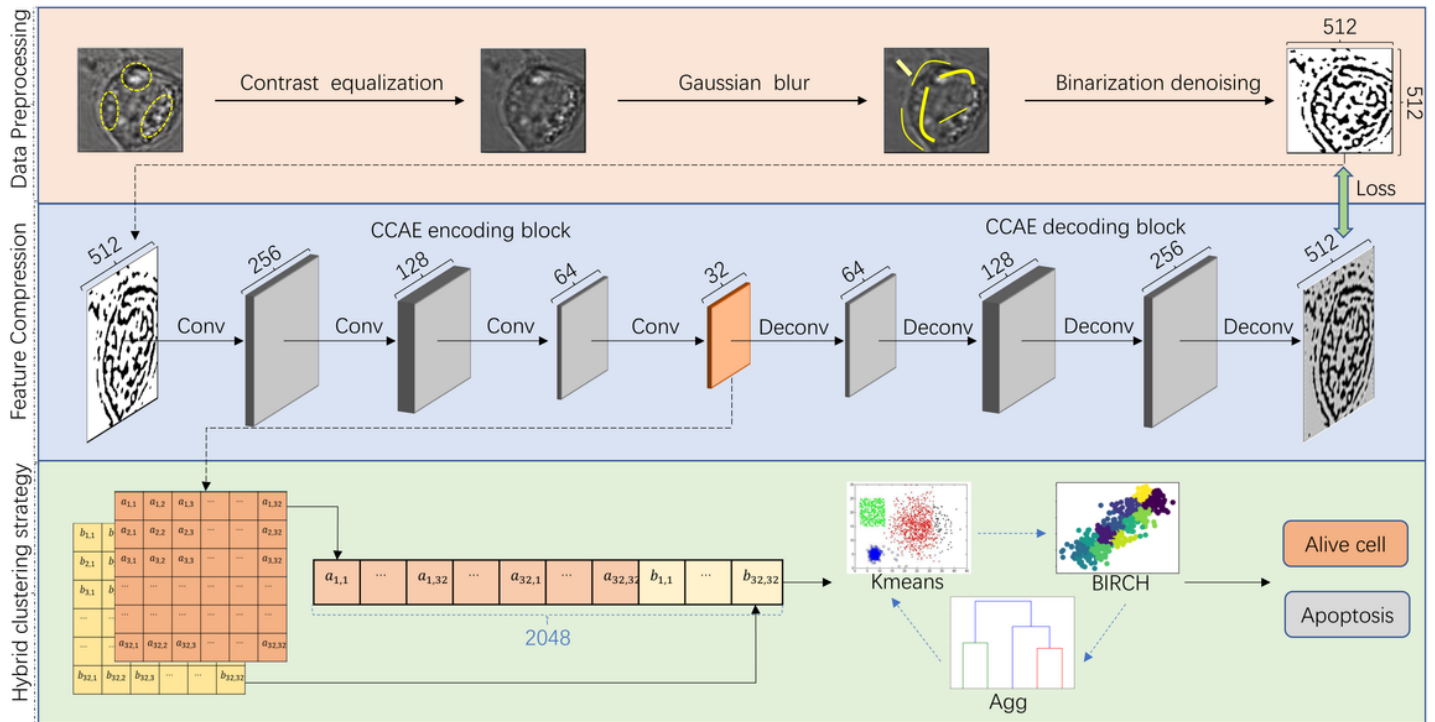


Figure 2

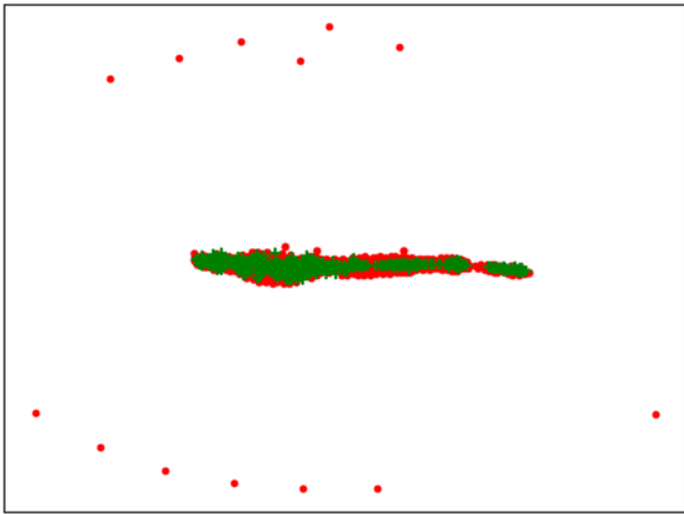
The structure and parameter setting of CCAE. The CCAE only uses convolutional and deconvolution layers, instead of filtering the features using a pooling or upsampling method that may lose important information. This shows several steps within the CCAE, starting from an image. The cropped cells are resized to 512×512 pixels before submitting to CCAE. To extract information from these images, several blocks of convolution are applied after the input has been resized. By applying a convolution matrix on

top of the original image, the image is recalculated and enhanced based on the values of surrounding pixels. Followed by a few deconvolution layers, these layers will reconstruct image from  $32 \times 32$  pixels to  $512 \times 512$  pixels. The latent vector parameter was sampled from a Gaussian distribution with a mean value of 0. Lastly, input image, reconstructed image and latent vector all used as parameters of triplet loss and help CCAE learn more effective features of cells.

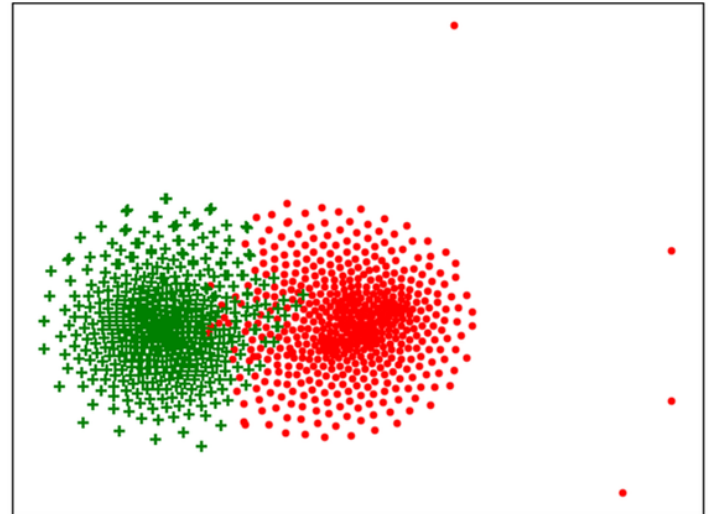


**Figure 3**

Flow chart of main method. The first step of the model is cell image preprocessing, including Gaussian blur, adaptive binarization and denoising. Afterwards, processed images are resized to  $512 \times 512$  pixels before submitting to CCAE model. After all the convolution and deconvolution steps, the encoding data (at the end of the convolutions of size  $32 \times 32 \times 2$ ), is flattened to be used in the following step of the model. The hybrid clustering algorithm, including the Euclidean distance k-mean algorithm, the agglomerative (AGG) algorithm, and the balanced iterative reducing and clustering using hierarchies (BIRCH) algorithm, was applied to get clustering results.



CCAЕ without preprocessing



CCAЕ with preprocessing

**Figure 4**

The effect of preprocessing on data compression. In order to demonstrate the effectiveness of preprocessing, the t-SNE method is used to visualize the feature maps of CCAE using the same dataset. According to the map of CCAE with preprocessing, samples are almost separated in two dimensions. In contrast, using CCAE without preprocessing features makes it challenging to distinguish two classes. This result suggests that the features extracted with preprocessing are more discriminative than those extracted without preprocessing and help CCAE learn more effective features of cells.