Mathematical Biosciences and Engineering

MBE, 20(5): 9423–9442.
DOI: 10.3934/mbe.2023414
Received: 23 December 2022
Revised: 06 March 2023
Accepted: 10 March 2023
Published: 20 March 2023

http://www.aimspress.com/journal/MBE

Research article

Classification and recognition of milk somatic cell images based on PolyLoss and PCAM-Reset50

Jie Bai[1,2], Heru Xue[1,2,*], Xinhua Jiang[1,2] and Yanqing Zhou[1,2]

[1] College of Computer and Information Engineering Inner Mongolia Agricultural University, Hohhot 010018, China
[2] Inner Mongolia Autonomous Region Key Laboratory of Big Data Research and Application of Agriculture and Animal Husbandry, Hohhot 010018, China

* **Correspondence:** Email: xuehr@126.com.

Abstract: Somatic cell count (SCC) is a fundamental approach for determining the quality of cattle and bovine milk. So far, different classification and recognition methods have been proposed, all with certain limitations. In this study, we introduced a new deep learning tool, i.e., an improved ResNet50 model constructed based on the residual network and fused with the position attention module and channel attention module to extract the feature information more effectively. In this paper, macrophages, lymphocytes, epithelial cells, and neutrophils were assessed. An image dataset for milk somatic cells was constructed by preprocessing to increase the diversity of samples. PolyLoss was selected as the loss function to solve the unbalanced category samples and difficult sample mining. The Adam optimization algorithm was used to update the gradient, while Warm-up was used to warm up the learning rate to alleviate the overfitting caused by small sample data sets and improve the model's generalization ability. The experimental results showed that the classification accuracy, precision rate, recall rate, and comprehensive evaluation index F value of the proposed model reached 97%, 94.5%, 90.75%, and 92.25%, respectively, indicating that the proposed model could effectively classify the milk somatic cell images, showing a better classification performance than five previous models (i.e., ResNet50, ResNet18, ResNet34, AlexNet andMobileNetv2). The accuracies of the ResNet18, ResNet34, ResNet50, AlexNet, MobileNetv2, and the new model were 95%, 93%, 93%, 56%, 37%, and 97%, respectively. In addition, the comprehensive evaluation index F1 showed the best effect, fully verifying the effectiveness of the proposed method in this paper. The proposed method overcame the limitations of image preprocessing and manual feature extraction by traditional machine learning methods and the limitations of manual feature selection, improving the classification accuracy

and showing a strong generalization ability.

Keywords: milk somatic cells; image processing; residual network; attention mechanism; loss function

1. Introduction

Mastitis is the most common infectious disease in dairy cows, causing huge economic losses to dairy farmers and milk processing companies [1,2]. The cells found in healthy milk comprise 75%–85% of white blood cells (macrophages, neutrophils, and lymphocytes) and 15%–25% of epithelial cells. During mastitis, the count of white blood cells can reach 99%, a useful biomarker to indicate infection [3,4]. Yet, the severity of cow mastitis infection and the counts of various cells may sometimes differ [3,4]. Somatic cell count (SCC) is a key method used in the dairy industry to determine the quality of cattle and bovine milk [5]. SCC measurement can be divided into direct and indirect methods. A direct approach, which relies on manual counting under the microscope or calculating and analyzing the number of somatic cells using computerized image processing techniques from the microscopic images of breast samples, is considered the most accurate method. However, the manual counting method is susceptible to the subjective factors of the experimenter, is time-consuming and less automated, and may lead to counting errors [6,7]. Indirect methods include the California cell assay and the Wisconsin mastitis test [8]. The most popular indirect approaches are cell analyzers (e.g., MoFlo XDP ultra-fast flow cytometer from Beckman Coulter (USA), Nucleo Counter NC-3000 cell analyzer from ChemoMetecA/S (Denmark), and Fossomatic™ 7 DC somatic cell analyzer from FOSS (Denmark)); yet, these methods are expensive and subject to slightly less correlated results and have poor accuracy.

Image processing technology has been increasingly used for classifying and identifying different cell types, mainly by collecting the color images of stained cells using a microscope [9]. Conventional recognition methods are mainly based on the following steps: obtaining the color images of cells through machine vision technology, extracting the cell features, and building models for cell classification and recognition using related machine learning methods [9]. Khan et al. [10] classified malignant and benign cells using a support vector machine (SVM) by extracting the texture information from breast cells. Okmen et al. [11] used the k-Nearest Neighbor (KNN) method to classify renal tumor cells, showing an accuracy of 83.8%. Moreover, Mishra et al. [12] extracted multiple features and classified lymphocytes using the random forest classifier. Gao et al. [13] proposed a rapid and accurate method for detecting mastitis in cattle by using the two-way two-dimensional principal component analysis. In addition, Gao et al. [14,15] proposed a Relief-F algorithm to extract the features of milk somatic cells for classification. Besides, Zhang et al. [16] proposed a recognition algorithm based on random forest, with an accuracy of 96%. Machine learning is also used for microorganism identification and classification [17–20].

Furthermore, over the last few years, novel image processing-based cell counting methods have been proposed; they combine the diagnostic experience of pathologists with the advantages of rapid processing and accurate calculation of computers to achieve the counting and classification of somatic cells, reducing the influence of subjective factors, and providing a convenient and accurate detection method. Liang et al. [21] proposed a set containing convolutional neural network (CNN) (using Xception) and recurrent neural network (RNN) (using long and short-term memory (LSTM)) features, with a cell classification accuracy of 90.79%. Moreover, Bani-Hani et al. [22] proposed applying CNN

and genetic algorithm (GA) for image classification of four types of blood cells, showing an accuracy of 91.00%. Habibzadeh et al. [23] used a pre-trained model, ResNet and InceptionNet, to identify white blood cells. Acevedo et al. [24] upgraded the structure of the Inception-v3 network model to classify and recognize eight types of blood cell images, reaching an accuracy of 90%. Malkawi et al. [25] extracted the features of leukocytes through the visual geometry group (VGG) network and then adopted the SVM model for classification and recognition, reaching an accuracy of 98.7%. Ghosh & Kundu [26] combined RNN and CNN to establish a multilayer network model and classify different blood cells, achieving an accuracy of 87.29%. AlexNet is the name of a convolutional neural network (CNN) architecture; AlexNet contains eight layers; the first five are convolutional layers, some of them followed by max-pooling layers, and the last three are fully connected layers [27]. MobileNetV2 is a convolutional neural network that is 53 layers deep and is a very effective feature extractor for object detection and segmentation [28].

Over recent years, an increasing number of studies started employing deep learning methods for microorganism detection [17–19]. The recognition model constructed by the deep learning algorithms is portable and has the advantages of automatic learning and extracting shallow and deep features [29]. Therefore, research and construction of a recognition model based on deep learning can reduce the difficulty of manual feature selection and present good mobility, suitable for milk somatic cell images with complex backgrounds [30]. In this study, we introduced PolyLoss and product of cross-attention matrices (PCAM)-ResNet50, a model based on the improved structure of ResNet50, and applied it to bovine milk somatic cell datasets to solve the problem of multiple classifications in which the dataset samples themselves are difficult to obtain, and the data are unevenly distributed. In consideration of the characteristics of milk cells, traditional data enhancement methods were used to expand the data set, increase its diversity, and reduce its overfitting. The PCAM module was introduced into the algorithm to make the network learn with the target, while PolyLoss was selected as the loss function to solve the difficult acquisition of samples and imbalance classification to improve the detection and classification accuracy. This new, improved method (i.e., PCAM-ResNet50 combined with PolyLoss) introduces a position attention module and a channel attention module into the classical ResNet50 network structure, respectively, which allows for more efficient and accurate extraction of cell features. Hence, the identification of the milk cells could be more accurate, providing more accurate identification of milk quality.

The workflow of the proposed method in this paper is shown in Figure 1.

The main contributions of this paper are as follows. 1) To our knowledge, this is the first study that uses deep learning techniques to classify images of bovine milk somatic cells. 2) This study proposes a new method (PCAM-ResNet50) that introduces a position attention module and a channel attention module into the classical ResNet50 network structure, respectively, to make it more effective in extracting the features of various types of cells. 3) The loss function of the original ResNet50 structure is improved, and PolyLoss was adopted to compensate for the poor classification effect caused by the lack of sample size of bovine milk cells and the unbalanced data set. 4) The performance of the proposed method is better than other methods, and the classification accuracy reaches 97%.

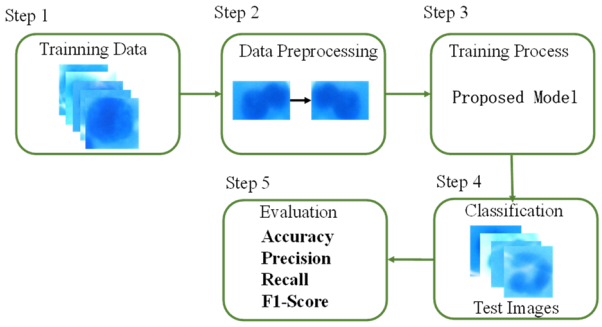
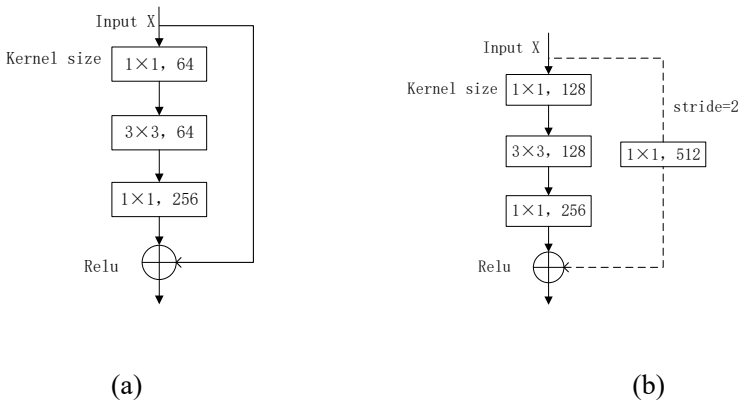


Figure 1. Study workflow. Step 1: a single bovine milk somatic cell image is used as a training sample and the details of the dataset. Step 2: the dataset is preprocessed and enhanced by geometric rotation and other methods. Step 3: the training data are fed into the models proposed in this paper (PolyLoss and PCAM-ResNeet50) for training. Step 4: the unseen test images are provided to the model in this paper for classification. Step 5: the performance of the proposed model is evaluated by calculating the accuracy, precision, recall, and F1 values.

2. Materials and methods

2.1. ResNet model

In Figure 2, Relu is input as the activation function, and the feature image is x. First, the feature image is reduced by checking the convolution of 1 × 1, doing a 3 × 3 convolution operation, and restoring the dimension by convolution 1 × 1. The solid line in Figure 2(a) and the dashed line in Figure 2(b) represent the input downsampling function. Figure 2(a),(b) illustrate the residual block without changing the size and that by adding scale, respectively. The output feature image is half of the original height and width. After such a jump connection, the shallow features are identically mapped to the deep layer. The deep gradient can be directly transmitted back to the shallow layer during the back propagation, which solves network degradation and improves performance.



(a) (b)

Figure 2. ResNet residual module [31]. (a) Residual structure without changing scale. (b) Residual structure with added scale.

2.2. Loss function

The loss function here refers to the difference between the predicted value and the true value of the output after the training data are input into the model. In the classical ResNet50 network, the cross-entropy of each sample is averagely weighted to obtain the final value of the loss function, shown in Eq (1):

$$L_{CE} = -\frac{1}{N}\sum_{i=1}^{m}[y_i \, ln \, \hat{y}_i + (1 - y_i) \, ln(1 - \hat{y}_i)] \tag{1}$$

where N and m present the number of samples and categories, respectively, and y_i and $\hat{y}_i$ stand for the true value and the predicted value of class *i*, respectively.

2.3. Model refinement

2.3.1. Attention mechanism

Milk has large amounts of fat, protein, and cellular debris, regarded as image backgrounds [32]. When cells are stained, the contrast between cells and the background is affected due to uneven illumination. The attention mechanism [33] was introduced in the residual model to reduce the influences of the above factors and make the model more targeted, improving recognition accuracy. The attention mechanism makes the model focus on the important information in the image [34]. The position attention mechanism (PAM) refers to the spatial dependence on different positions on the feature map [35]. The channel attention mechanism (CAM) and soft attention were used here. At a particular position, the feature is updated by a weighted sum at all positions, where the weights are determined by the feature similarity between the two positions. PAM encodes a wide range of contextual information into local features to improve its representation power. Figure 3 illustrates the structure of PAM.

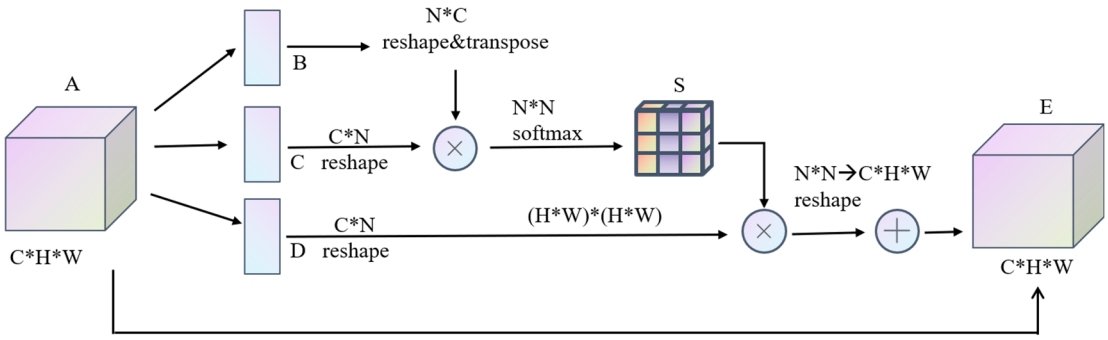


Figure 3. Position attention mechanism.

The calculation process of Figure 3 was as follows:

I. Three feature maps $\{B, C, D\} \in R^{C\times H\times W}$ were obtained after passing the input feature map $A \in R^{C\times H\times W}$ through three convolutional layers.

II. The transposes of the matrices C and B were multiplied to obtain the similarity matrix of the pixels within A, and the spatial attention map $S \in R^{N\times N}$ was obtained by Softmax with the following Eq (2).

$$S_{ji} = \frac{exp(B_i C_j)}{\sum_{i=1}^{N} exp(B_i C_j)} \tag{2}$$

where $N = H \times W$ represents the number of pixels in the spatial range, and S_{ji} is the relationship of the i th and j th positions. The more similar the characteristic representations of the two positions are, the more relevant they are.

III. Matrix multiplication was performed by transposing matrices S and D, and the result was restored to the original input feature size $R^{C\times H\times W}$. The final output result $E \in R^{C\times H\times W}$ was obtained by summing it with the original features element by element through the scale coefficient α, as shown in Eq (3).

$$E_j = \alpha \sum_{i=1}^{N} (S_{ji} D_i) + A_j \tag{3}$$

where α is initialized to 0 and gradually learns to assign more weights, D_i and A_j are the element of D and A, respectively. The above equations suggest that the feature E obtained at each position was the weighted sum of the features with the original features at all positions.

For the CAM [35], advanced features of each channel could be regarded as category-specific information extraction of the dependency in the channel dimension. All channel features were weighted, and each channel feature was updated to improve classification accuracy (Figure 4). In this paper, the correlation of channels was modeled by using the spatial information of all relevant positions.

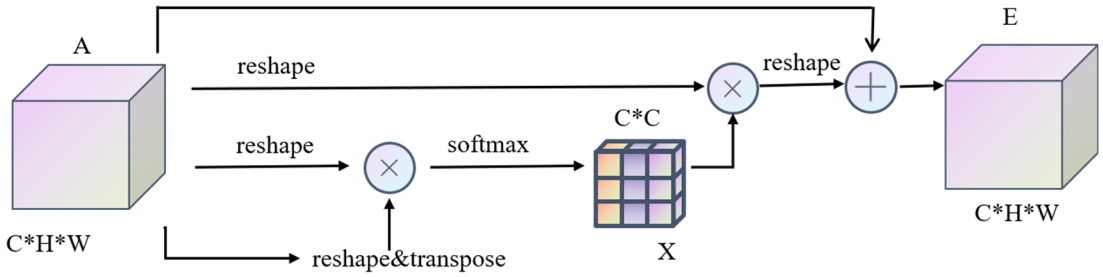


Figure 4. Channel attention mechanism.

As demonstrated in Figure 4, the channel attention map $A \in R^{C\times H\times W}$ was calculated directly from the input feature map $X \in R^{C\times C}$, as follows:

1. The input feature graph $A \in R^{C\times H\times W}$ was performed the reshape operation, which becomes$R^{C\times N}$, where $N = H \times W$.
2. The transition matrixes of $R^{C\times N}$ and $R^{C\times N}$ were multiplied to obtain the channel attention matrix $X \in R^{C\times C}$ by Softmax, as follows:

$$x_{ji} = \frac{exp(A_i \cdot A_j)}{\sum_{i=1}^{C} exp(A_i \cdot A_j)} \tag{4}$$

where, x_{ji} represents the effect of the i th channel on the j th channel.

3. The matrix $R^{C\times N}$ was obtained by multiplying the matrix X with A. After it was multiplied by the scale factor β, the element-level summing operation with A was performed to obtain the final output result $E \in R^{C\times H\times N}$. The specific calculation is shown in Eq (5):

$$E_j = \beta \sum_{i=1}^{C} (x_{ji} \cdot A_i) + A_j \tag{5}$$

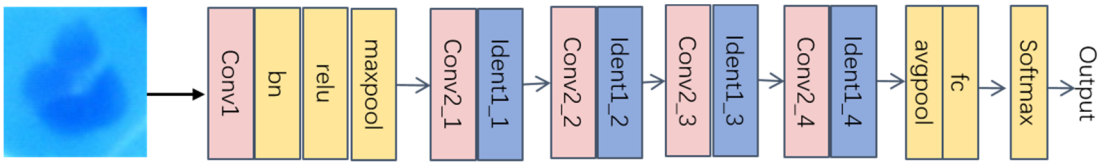
2.3.2. The PolyLoss function

Lin et al. [36] proposed to adopt the focus loss function to avoid the overfitting caused by the cross-entropy loss function, but the loss function is not the most effective on some unbalanced data sets. The PolyLoss function [37] was selected for the improved ResNet50 model in this paper to solve the poor classification effect caused by insufficient cells and class imbalance. PolyLoss was used to decompose the commonly used classification loss functions (such as cross-entropy loss function and focal loss function) into a series of weighted polynomial bases through Taylor expansion, as shown in Eq (6):

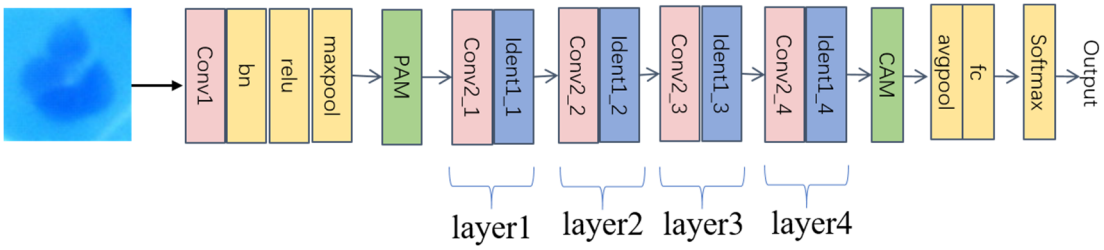
$$L_{Poly-N} = -log(P_t) + \sum_{j=1}^{N} \varepsilon_j (1 - P_t)^{\gamma} \tag{6}$$

where $\varepsilon_j > 0$ is the polynomial coefficient, $\gamma \in [-1/\gamma, \infty]$ $j \in [-1/j, \infty]$ is the perturbation term, and P_t represents the probability of the target-label prediction.

2.3.3. Resedual block model based on PCAM-ResNet50



(a)



(b)

Figure 5. Model structure diagram. (a) Classical ResNet50 model. (b) The PCAM-ResNet50 model.

The deep residual network model used in this paper was based on the classical ResNet50 model (Figure 5(a)) realized by integrating the PAM and CAM, forming a PCAM-ResNet50 model, as shown in Figure 5(b). The figure reveals that the input image sizes of the two model structures were the same, which was a color milk somatic cell image in 224 × 224 × 3. Moreover, the model structure included the conv layer, bn layer, maxpool layer, avgpool layer, fc layer, and Softmax classification layer. Where Conv1 was the convolution layer, Conv2_x (x = 1, 2, 3, 4) was the residual block with scale added, and there was only one residual block. Besides, Ident1_x (x = 1, 2, 3, 4) presented the residual block without changing the size, with 2, 3, 5, and 2 residual blocks, respectively. The improved part is marked in green in Figure 5(b). The PAM and CAM were added after the Max pooling layer and the last residual block, respectively. It aimed to ensure that the effective milk somatic cell image information could be transmitted to the deep layer, and the feature weight of the cell area was increased. The interference information, such as contrast and brightness, was reduced, and the impurity debris in the image background was filtered out to improve the accuracy of automatic recognition.

The network structure contained five convolutional blocks. The first was a 7 × 7 convolutional layer; the remaining four were residual blocks. Each residual block comprised three convolutional layers, with 1 + 3 × (3 + 4 + 6 + 3) = 49 convolutional layers. The specific parameter settings for the PCAM-ResNet50 are listed in Table 1.

Table 1. The parameter settings for the PCAM-ResNet50.

Layer Name	Kernel Size	Stride	Channels	Output Size
Conv1	7 × 7	2	64	112 × 112
maxpool	3 × 3	2	64	56 × 56
PAM	1 × 1	1	64	56 × 56
layer1	1 × 1	1	256	56 × 56
layer2	1 × 1	2	512	28 × 28
layer3	1 × 1	2	1024	14 × 14
layer4	1 × 1	2	2048	7 × 7
CAM			2048	7 × 7
avgpool			2048	1 × 1

The flow of the image classification and recognition method under the improved ResNet 50 model is shown in Figure 6.

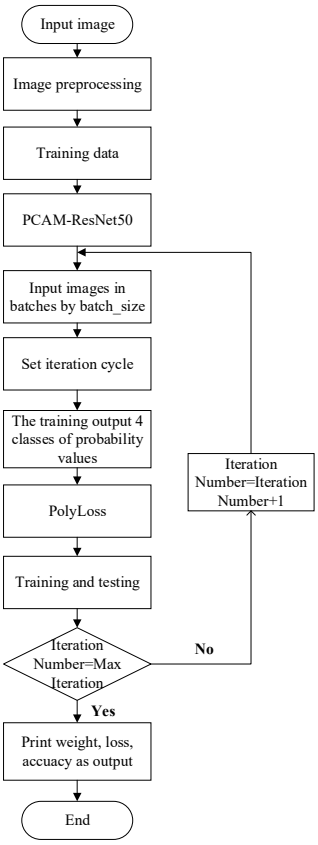


Figure 6. Flow chart of the algorithm proposed in this paper.

3. Results and discussion

3.1. Model training environment

This experiment was carried out using the Windows 10 operating system (central processing unit: Intel (R) Core (TM) i5-12400F CPU 2.50 GHz, running memory of 32 G; graphics card: NVIDIA GeForce RTX 2060). The open-source framework for Pytorch deep learning was used to build the deep neural network. CUDA11.6 was used to accelerate the network training. The programming was realized with Pycharm and Python3.8.5.

3.2. Experimental data set and preprocessing

The dataset used in this experiment was obtained from the Basic Veterinary Laboratory of Inner Mongolia Agricultural University. It was magnified 400× using an OLYMPUS BX51 optical microscope, with a resolution of 2048 pixels × 1536 pixels, and it contained 57 original large-field images (as shown in Figure 7), which were segmented using the Kmeans algorithm [38]. After extraction, a pathologist screened and manually classified the cells into four types. There were 168 macrophages (MΦ), 68 epithelial cells (EPI), 1209 neutrophils (NG), and 153 lymphocytes (LYP), corresponding to the numbers 1, 2, 3, and 4 in Figure 7, respectively. A total of 1598 images were initially analyzed. In order to enhance the generalization ability of the model and reduce overfitting,

the sample size was expanded by random rotation, horizontal flip, vertical flip, and random grayscale. Hence, 7990 images were finally obtained and randomly divided into a training set and a test set using a ratio of 8:2. The data samples are shown in Figure 8.

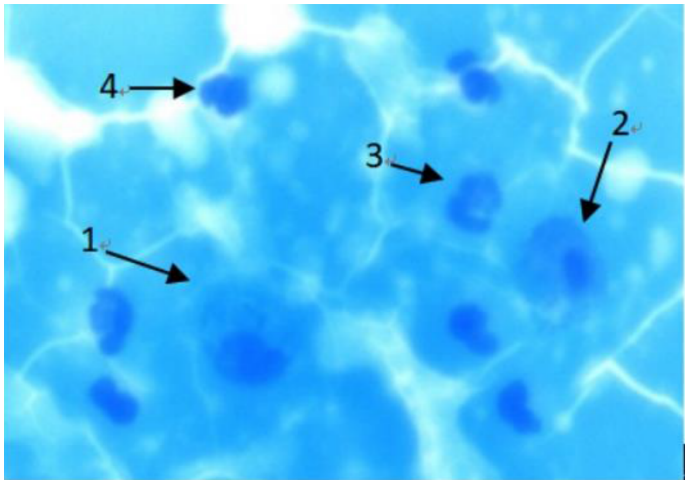


Figure 7. Original large-field image.

(a) Original EPI | (b) Random Rotation | (c) Turn over horizontally | (d) Turn over vertically | (e) Random grayscale

Figure 8. Example image of sample data.

3.3. Experimental parameter setting

In order to ensure the model was in the optimal state, the parameters were set after several experimental verifications as follows: the iteration times of the network training epoch were set to 200, the batch size was set to 16, and the Adam optimizer [39] was adopted (its initial learning rate was 0.001). However, the Warm-up learning rate was adopted to optimize the Adam algorithm. The Warm-up stage of network training was performed by gradually increasing the low learning rate. The initial learning rate was used for training when the model was relatively stable. At this time, the learning rate gradually decreased, accelerating the convergence and improving the effect.

3.4. Model evaluation method

The classification results were shown through the confusion matrix to evaluate the effectiveness of the identification method in this paper. The accuracy (A), accuracy (P), recall rate (R), harmonic average (F1) of recall rate, and macro average F1 (F1-macro) were adopted for a comprehensive evaluation. They could be calculated with the following equations:

$$A = \frac{TP+TN}{TP+FP+TN+FN} \tag{7}$$

$$P = \frac{TP}{TP+FP} \tag{8}$$

$$R = \frac{TP}{TP+FN} \tag{9}$$

$$F1 = \frac{2PR}{P+R} \tag{10}$$

$$F1 - macro = \frac{1}{n}\sum_{i-1}^{n} F1^{(i)} \tag{11}$$

In the above equations, *TP* represents the true type of cell sample, i.e., the number identified by the model as the correct type; *FN* indicates that the cell sample was the true type, i.e., the number of types identified as wrong by the model; *FP* represents the wrong type of cell sample, i.e., the number of the correct type in the recognition model; *TN* is the cell sample is the error type, i.e., the recognition model is also the number of error types. In addition, *n* refers to the number of cell types.

A relational diagram of evaluation indicators was constructed (Figure 9) to elucidate the relationship among *TP*, *FN*, *FP*, and *TN*, where *P* and *N* represent the correct and wrong samples, respectively.

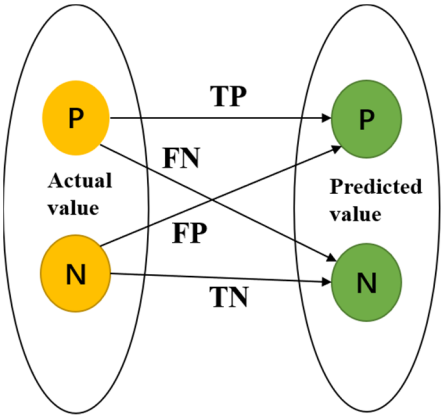


Figure 9. Relationship diagram of evaluation indicators.

In the experiment, the receiver operating characteristic (ROC) curve was used to evaluate the model performance. The abscissa is the false positive rate (FPR), i.e., the proportion of all the predicted samples of the correct type but the wrong type in all the wrong samples, as shown in Eq (12). The ordinate marks the true positive rate (TPR), i.e., the proportion of all the predicted samples of the correct type in all the samples of the correct type, as shown in Eq (13). The effect of the model was evaluated by calculating the area under the ROC curve (AUC) and the coordinate axis. The larger the value, the better the classification effect of the model.

$$FPR = \frac{FP}{FP+TN} \tag{12}$$

$$TPR = \frac{TP}{TP+FN} \tag{13}$$

3.5. Result analysis

3.5.1. Loss function validity analysis

In order to verify the effectiveness of the Polyloss function for the classification of unbalanced milk somatic cell samples, the model proposed in this paper was compared with the classic ResNet50. The experimental results are shown in Table 2. The macro average F1 value of the proposed algorithm was 8% higher than that of the classic ResNet50 model, indicating that the model pays more attention to the sample imbalance after introducing the Polyloss function, thus improving the generalization ability of the model.

Table 2. Comparative analysis of the effectiveness of the Polyloss function.

Method	F1-macro/%
ResNet50	84
New Model	92

Meanwhile, the loss functions of the new model, ResNet18, ResNet34, and ResNet50 model training, were drawn, as shown in Figure 10. With the increase in training times, the loss value of the model in this paper greatly decreased, so the model was relatively stable. During training, the loss value of the model in this paper was 1/7 that of other methods. Besides, the curve tends to converge as the number of iterations increases.

3.5.2. Analysis of ablation experiment results

The ResNet50 model, the ResNet50 model using the PolyLoss function, and the new model were applied to the test set to verify the overall impacts of each module. Table 3 shows that the accuracy of the ResNet50 network was 93%. It decreased by 1% after the loss function was replaced with the PolyLoss, while it increased by 4% after applying the proposed algorithm. Such data suggest that the new method can direct the network to focus more on the useful feature information in various cell images. Thus, the network can enhance the expression ability of milk somatic cell features and more accurately classify all types of milk cells.

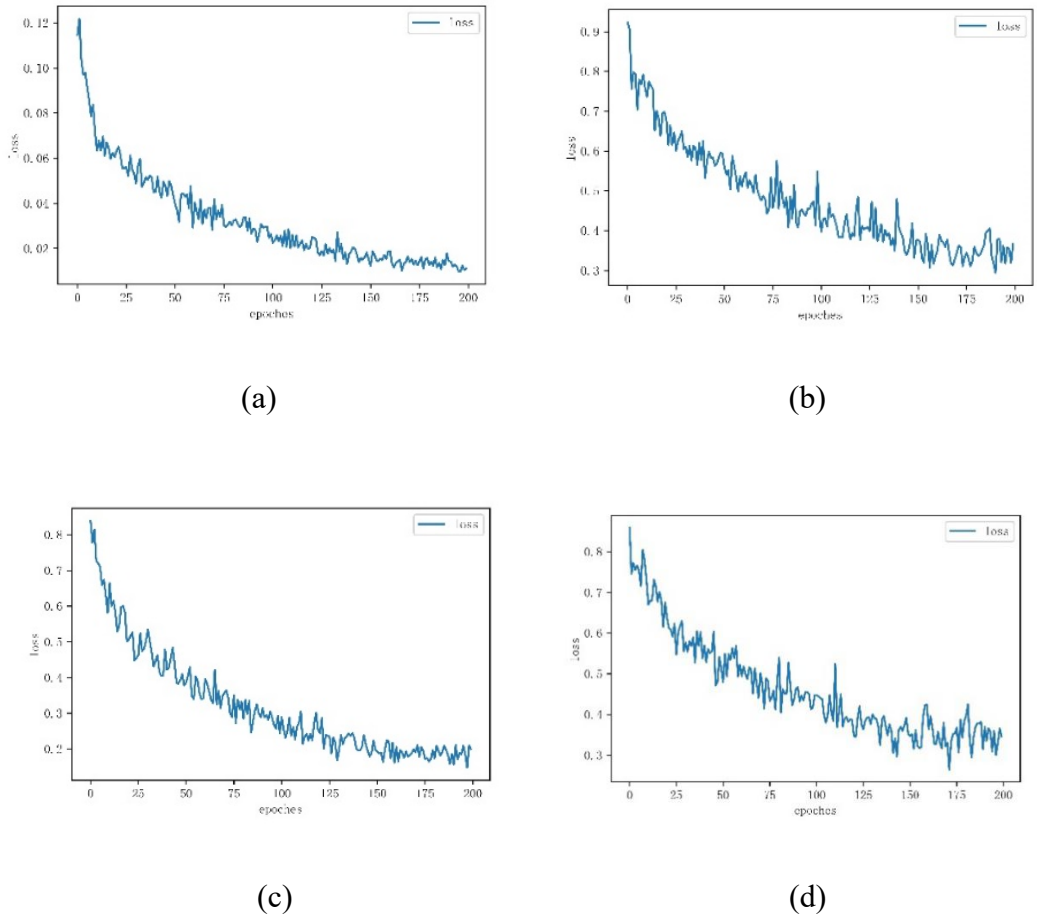


Figure 10. Training loss curves of different models. (a) The new model in this paper. (b) ResNet50. (c) ResNet18. (d) ResNet34.

Table 3. Ablation experiment results.

Model	PolyLoss	Attention mechanism	Accuracy
ResNet50	×	×	93%
ResNet50-PolyLoss	√	×	92%
New model	√	√	97%

3.5.3. Comparison of overall recognition effect of models

The overall recognition effect of the proposed algorithm was compared to ResNet18, ResNet34, ResNet50, AlexNet [27], and MobileNetV2 [28] models to verify the superiority of the proposed algorithm. As shown in Table 4, the accuracy, precision rate, recall rate, and F1 value of the new proposed algorithm were 97%, 94.5%, 90.75%, and 92.25%, respectively, and were higher than those of the other five methods, indicating that the proposed algorithm is more suitable for milk somatic cell image recognition. Meanwhile, the ROC curves of the six recognition models generated are shown in Figure 11.

Table 4. Comparison of recognition effects of different models.

Classification model	Overall recognition effect/%			
	Accuracy	Precision	Recall	F1-score
ResNet18	95	91.5	87	89.25
ResNet34	93	86.25	81.25	83.5
ResNet50	93	87.75	81.25	83.5
New model	97	94.5	90.75	92.25
AlexNet	56	26	24.75	25
MobileNetV2	37	21	21	17

3.5.4. The recognition effect of different models on different types of cells.

In order to verify the effectiveness of the proposed algorithm for milk somatic cell classification, it was compared and evaluated with the other three models in terms of precision rate, recall rate, and F1 comprehensive index. The comparison results are shown in Tables 5–7. In addition, after 1598 samples are tested, the confusion matrix of the four types of cells is shown in Figure 12.

Table 5. Comparison of precision rates of different classification models.

Classification model	Precision rate of each class/%			
	MΦ	LYP	EPI	NG
ResNet18	89	89	91	97
ResNet34	88	81	80	96
ResNet50	87	79	88	97
This paper	93	93	93	99

Table 6. Comparison of recall rates of different classification models.

Classification model	Recall rate of each class/%			
	MΦ	LYP	EPI	NG
ResNet18	79	91	79	99
ResNet34	71	86	70	98
ResNet50	71	92	64	98
This paper	92	93	78	100

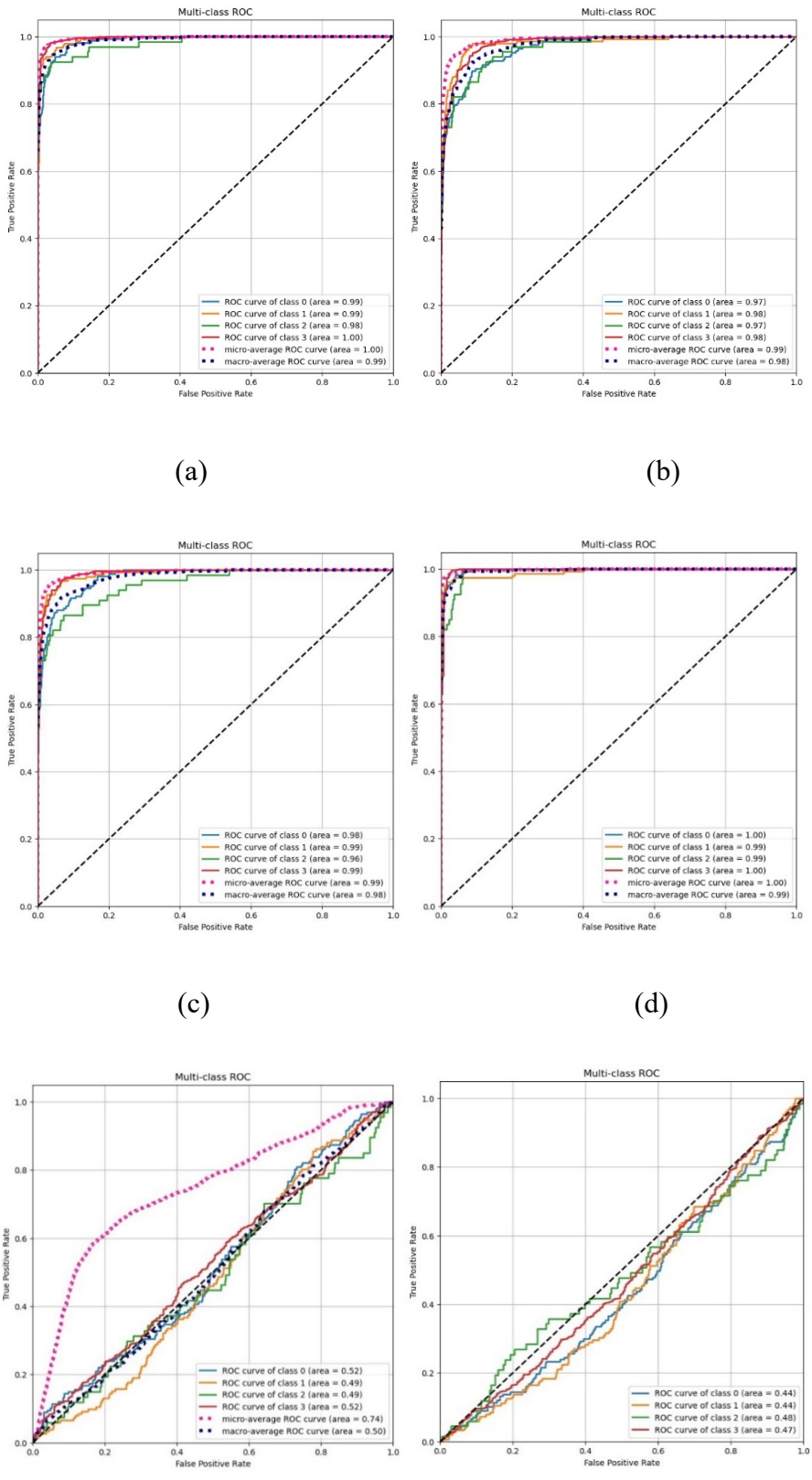
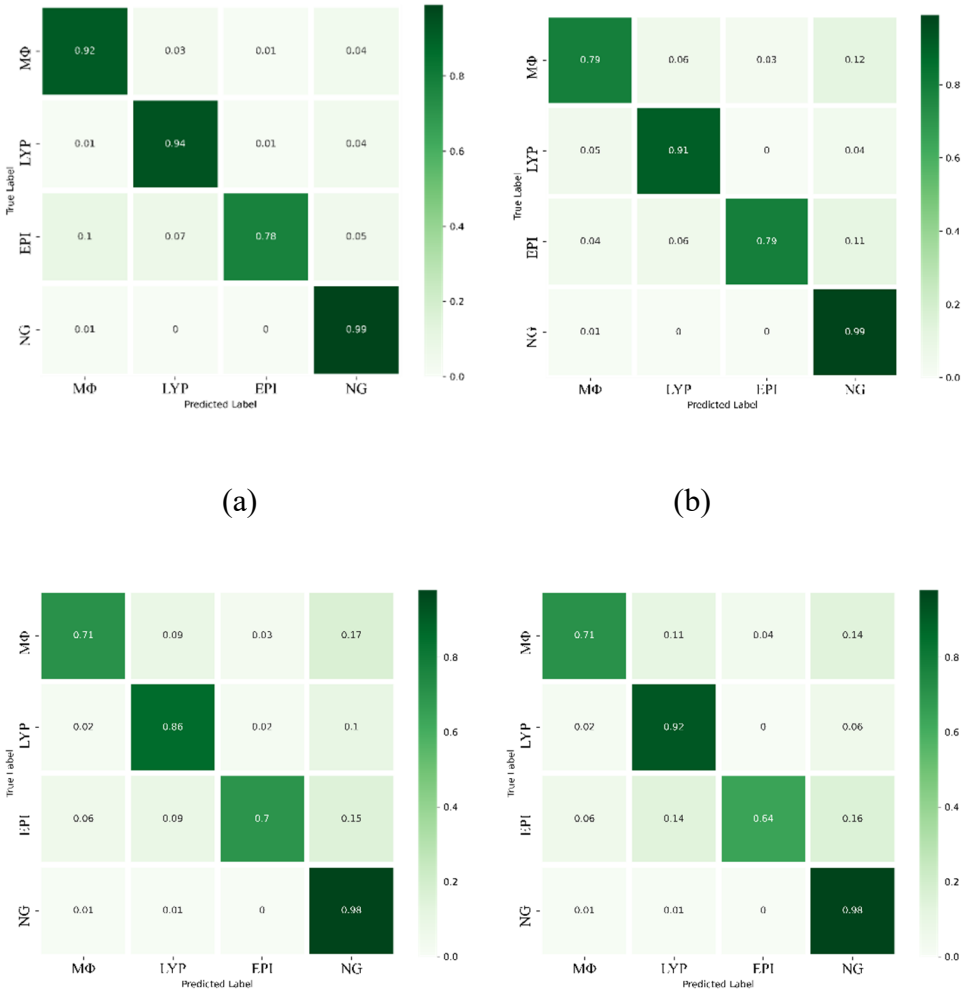


Figure 11. ROC curves of different classification models. (a) ResNet18. (b) ResNet34. (c) ResNet50. (d) New modeI(e) AlexNet. (f) MobileNetV2.

Table 7. Comparison of F1-score of different classification models.

Classification model	F1-score of each class/%			
	МΦ	LYP	EPI	NG
ResNet18	84	90	85	98
ResNet34	79	83	75	97
ResNet50	78	85	74	97
This paper	92	93	85	99



(a) (b)

(c) (d)

Figure 12. Confusion matrix results of different models. (a) New model. (b) ResNet18. (c) ResNet34. (d) ResNet50.

4. Conclusions

The new method (i.e., PCAM-ResNet50) introduces a position attention module and a channel attention module into the classical ResNet50 network structure, respectively, making the older tool more effective in extracting the features of various types of cells. In addition, the loss function of the original ResNet50 structure was improved, and the polynomial-based combination loss function (PolyLoss) was adopted to make up for the poor classification effect caused by the lack of sample size of bovine milk cells and the unbalanced data set. The performance of the new method was superior to other methods, and the classification accuracy reached 97%. Hence, the identification of the milk cells could be more accurate, providing more accurate identification of milk quality. The proposed method in this paper also has the potential to be used in some other research fields about AI, such as image segmentation [40], medical image analysis [41,42], feature extraction [43], and video analysis [44].

Acknowledgments

This study was supported by the Inner Mongolia Autonomous Region Higher Education Scientific Research Project (#NJZY21486), the National Natural Science Foundation of China (#31960494), the Inner Mongolia Autonomous Region Science and Technology Project (#2020GG0169), and the Inner Mongolia Agricultural University Basic Subject Scientific Research Funding Project (#JC2018001).

Conflict of interest

The authors declare there is no conflict of interest.

References

1. T. Halasa, K. Huijps, O. Østerås, H. Hogeveen, Economic effects of bovine mastitis and mastitis management: a review, *Vet. Q.*, **29** (2007), 18–31. https://doi.org/10.1080/01652176.2007.9695224
2. U. Geary, N. Lopez-Villalobos, N. Begley, F. Mccoy, B. O. Brien, L. O. Grady, Estimating the effect of mastitis on the profitability of irish dairy farms, *J. Dairy Sci.*, **95** (2012), 3662–3673. https://doi.org/10.3168/jds.2011-4863
3. D. Barrett, High somatic cell counts-a persistent problem, *Irish Vet. J.*, **55** (2002), 173–178.
4. H. M. Golder, A. Hodge, I. J. Lean, Effects of antibiotic dry-cow therapy and internal teat sealant on milk somatic cell counts and clinical and subclinical mastitis in early lactation, *J. Dairy Sci.*, **99** (2016), 7370–7380. https://doi.org/10.3168/jds.2016-11114
5. J. Hamann, Changes in milk somatic cell count with regard to the milking process and the milking frequen, *Bull. Int. Dairy Fed.*, **24** (2001), 5–6.
6. G. Leitner, Y. Lavon, Z. Matzrafi, O. Benun, D. Bezman, U. Merin, Somatic cell counts, chemical composition and coagulation properties of goat and sheep bulk tank milk, *Int. Dairy J.*, **58** (2016), 9–13. https://doi.org/10.1016/j.idairyj.2015.11.004
7. U. K. Sundekilde, N. A. Poulsen, L. B. Larsen, H. C. Bertram, Nuclear magnetic resonance metabonomics reveals strong association between milk metabolites and somatic cell count in bovine milk, *J. Dairy Sci.*, **96** (2013), 290–299. https://doi.org/10.3168/jds.2012-5819

8. J. S. Moon, H. C. Koo, Y. S. Joo, S. H. Jeon, D. S. Hur, C. I. Chung, Application of a new portable microscopic somatic cell counter with disposable plastic chip for milk analysis, *J. Dairy Sci.*, **90** (2007), 2253–2259. https://doi.org/10.3168/jds.2006-622
9. A. Awad, and M. Hassaballah, Image Feature Detectors and Descriptors, *Springer*, (2016).
10. S. U. Khan, N. Islam, Z. Jan, K. Haseeb, S. I. A. Shah, M. Hanif, A machine learning-based approach for the segmentation and classification of malignant cells in breast cytology images using gray level co-occurrence matrix (GLCM) and support vector machine (SVM), *Neural Comput. Appl.*, **34** (2022), 8365–8372. https://doi.org/10.1007/s00521-021-05697-1
11. H. B. ökmen, A. Guvenis, H. Uysal, Predicting the polybromo-1 (PBRM1) mutation of a clear cell renal cell carcinoma using computed tomography images and knn classification with random subspace, *Vibroeng. Procedia*, **26** (2019), 30–34. https://doi.org/10.21595/vp.2019.20931
12. S. Mishra, B. Majhi, P. K. Sa, L. Sharma, Gray level co-occurrence matrix and random forest based acute lymphoblastic leukemia detection, *Biomed. Signal Process. Control*, **33** (2017), 272–280. https://doi.org/10.1016/j.bspc.2016.11.021
13. X. J. Gao, H. R. Xue, X. Pan, X. H. Jiang, Y. Q. Zhou, X. L. Luo, Somatic cells recognition by application of Gabor feature-based $(2D)^2$PCA, *Int. J. Pattern Recognit Artif. Intell.*, **31** (2017), 1757009. https://doi.org/10.1142/S0218001417570099
14. X. J. Gao, H. R. Xue, X. Pan, X. L. Luo, Polymorphous bovine somatic cell recognition based on feature fusion, *Int. J. Pattern Recognit Artif. Intell.*, **34** (2020), 2050032. https://doi.org/10.1142/S0218001420500329
15. X. J. Gao, H. R. Xue, X. H. Jiang, Y. Q. Zhou, Recognition of somatic cells in bovine milk using fusion feature, *Int. J. Pattern Recognit Artif. Intell.*, **32** (2018), 1850021. https://doi.org/10.1142/S0218001418500210
16. X. L. Zhang, H. R. Xue, X. J. Gao, Y. Q. Zhou, Milk somatic cells recognition based on multi-feature fusion and random forest, *J. Inn. Mongolia Agric. Univ.*, **39** (2018), 87–92. https://doi.org/10.16853/j.cnki.1009-3575.2018.06.014
17. J. Zhang, C. Li, M. M. Rahaman, Y. Yao, P. Ma, J. Zhang, et al., A comprehensive survey with quantitative comparison of image analysis methods for microorganism biovolume measurements, *Arch. Comput. Methods Eng.*, **30** (2023), 639–673. https://doi.org/10.1007/s11831-022-09811-x
18. P. Ma, C. Li, M. Rahaman, Y. Yao, J. Zhang, S. Zou, et al., A state-of-the-art survey of object detection techniques in microorganism image analysis: from classical methods to deep learning approaches, *Artif. Intell. Rev.*, **56** (2023), 1627–1698. https://doi.org/10.1007/s10462-022-10209-1
19. J. Zhang, C. Li, Y. Yin, J. Zhang, M. Grzegorzek, Applications of artificial neural networks in microorganism image analysis: a comprehensive review from conventional multilayer perceptron to popular convolutional neural network and potential visual transformer, *Artif. Intell. Rev.*, **56** (2023), 1013–1070. https://doi.org/10.1007/s10462-022-10192-7
20. J. Zhang, C. Li, M. M. Rahaman, Y. Yao, P. Ma, J. Zhang, et al., A comprehensive review of image analysis methods for microorganism counting: from classical image processing to deep learning approaches, *Artif. Intell. Rev.*, **55** (2022), 2875–2944. https://doi.org/10.1007/s10462-021-10082-4
21. G. Liang, H. Hong, W. Xie, L. Zheng, Combining convolutional neural network with recursive neural network for blood cell image classification, *IEEE Access*, **6** (2018), 36188–36197. https://doi.org/10.1109/ACCESS.2018.2846685

22. D. Bani-Hani, N. Khan, F. Alsultan, S. Karanjkar, N. Nagarur, Classification of leucocytes using convolutional neural network optimized through genetic algorithm, in *Proceedings of the 7th Annual World Conference of the Society for Industrial and Systems Engineering*, (2018), 1–6.
23. M. Habibzadeh, M. Jannesari, Z. Rezaei, H. Baharvand, M. Totonchi, Automatic white blood cell classification using pre-trained deep learning models: resnet and inception, in *10th International Conference on Machine Vision (ICMV 2017)*, (2018), 274–281. https://doi.org/10.1117/12.2311282
24. A. Acevedo, S. Alférez, A. Merino, L. Puigví, J. Rodellar, Recognition of peripheral blood cell images using convolutional neural networks, *Comput. Methods Programs Biomed.*, **180** (2019), 105020. https://doi.org/10.1016/j.cmpb.2019.105020
25. A. Malkawi, R. Al-Assi, T. Salameh, H. Alquran, A. M. Alqudah, White blood cells classification using convolutional neural network hybrid system, in *2020 IEEE 5th Middle East and Africa Conference on Biomedical Engineering (MECBME)*, (2020), 1–5. https://doi.org/10.1109/MECBME47393.2020.9265154
26. I. Ghosh, S. Kundu, Combining neural network models for blood cell classification, *arXiv preprint*, 2021, arXiv:2101.03604. https://doi.org/10.48550/arXiv.2101.03604
27. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, et al., Imagenet large scale visual recognition challenge, *Int. J. Comput. Vision*, **115** (2015), 211–252. https://doi.org/10.1007/s11263-015-0816-y
28. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L. Chen, Mobilenetv2: in-verted residuals and linear bottlenecks, in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (2018), 4510–4520. https://doi.org/10.1109/CVPR.2018.00474
29. Y. Hu, D. Y. Luo, K. Hua, H. M. Lu, X. G. Zhang, Overview on deep learning, *CAAI Trans. Intell. Syst.*, **14** (2019), 1–19. https://doi.org/10.11992/tis.201808019
30. A. I. Awad, M. Hassaballah, *Deep Learning in Computer Vision*, CRC Press, 2021.
31. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2016), 770–778. https://doi.org/10.1109/CVPR.2016.90
32. X. J. Gao, *Research of Polymorphous Bovine Somatic Cell Recognition Based on Feature Fusion*, Ph.D thesis, Inner Mongolia Agricultural University in Huhhot, 2018.
33. Y. Li, S. Tong, T. Li, Composite adaptive fuzzy output feedback control design for uncertain nonlinear strict-feedback systems with input saturation, *IEEE Trans. Cybern.*, **45** (2015), 2299–2308. https://doi.org/10.1109/TCYB.2014.2370645
34. Z. F. Jiang, T. He, Y. L. Shi, X. Long, S. H. Yang, Remote sensing image classification based on convolutional block attention module and deep residual network, *Laser J.*, **43** (2022), 76–81. https://doi.org/10.14016/j.cnki.jgzz.2022.04.076
35. J. Fu, J. Liu, H. J. Tian, Y. Li, Y. J. Bao, Z. W. Fang, et al., Dual attention network for scene segmentation, in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, (2019), 3141–3149. https://doi.org/10.1109/CVPR.2019.00326
36. T. Lin, P. Goyal, R. Girshick, K. He, P. Dollar, Focal loss for dense object detection, in *2017 IEEE International Conference on Computer Vision (ICCV)*, (2017), 2999–3007. https://doi.org/10.1109/ICCV.2017.324

37. Z. Q. Leng, M. X. Tan, C. X. Liu, E. D. Cubuk, X. J. Shi, S. Y. Cheng, et al., Polyloss: a polynomial expansion perspective of classification loss functions, *arXiv preprint*, arXiv:2204.12511. https://doi.org/10.48550/arXiv.2204.12511
38. J. Bai, H. R. Xue, X. H. Jiang, Y. Q. Zhou, Recognition of bovine milk somatic cells based on multi-feature extraction and a gbdt-adaboost fusion model, *Math. Biosci. Eng.*, **19** (2022), 5850–5866. https://doi.org/10.3934/mbe.2022274
39. D. P. Kingma, J. Ba, Adam: a method for stochastic optimization, arXiv preprint, arXiv:1412.6980. https://doi.org/10.48550/arXiv.1412.6980
40. J. Zhang, C. Li, S. Kosov, M. Grzegorzek, K. Shirahama, T. Jiang, et al., LCU-Net: A novel low-cost U-Net for environmental microorganism image segmentation, *Pattern Recognit.*, **115** (2021), 107885. https://doi.org/10.1016/j.patcog.2021.107885
41. H. Chen, C. Li, X. Li, M. M. Rahaman, W. Hu, Y. Li, et al., IL-MCAM: An interactive learning and multi-channel attention mechanism-based weakly supervised colorectal histopathology image classification approach, *Comput. Biol. Med.*, **143** (2022), 105265. https://doi.org/10.1016/j.compbiomed.2022.105265
42. X. Li, C. Li, M. M. Rahaman, H. Sun, X. Li, J. Z. Wu, et al., A comprehensive review of computer-aided whole-slide image analysis: from datasets to feature extraction, segmentation, classification and detection approaches, *Artif. Intell. Rev.*, **35** (2022), 4809–4878. https://doi.org/10.1007/s10462-021-10121-0
43. F. Kulwa, C. Li, J. Zhang, K. Shirahama, S. Kosov, X. Zhao, et al., A new pairwise deep learning feature for environmental microorganism image analysis, *Environ. Sci. Pollut. Res.*, **29** (2022), 51909–51926. https://doi.org/10.1007/s11356-022-18849-0
44. A. Chen, C. Li, S. Zou, M. M. Rahaman, Y. Yao, H. Chen, et al., SVIA dataset: A new dataset of microscopic videos and images for computer-aided sperm analysis, *Biocybern. Biomed. Eng.*, **42** (2022), 204–214. https://doi.org/10.1016/j.bbe.2021.12.010

AIMS Press