

Research Article

Efficient Cervical Cell Lesion Recognition Method Based on Dual Path Network

Mingwei Qin ^{1,2}, Yongxiang Gao ^{1,2}, Haifeng Sun ³, Mingliang Gou ^{1,2}, Ning Zhou ⁴,
and Yuancheng Yao ^{1,2}

¹Sichuan Key Laboratory of Special Environment Robotics, Mianyang, Sichuan 621010, China

²School of Information Engineering, Southwest University of Science and Technology, Mianyang, Sichuan 621010, China

³School of Computer Science and Technology, Southwest University of Science and Technology, Mianyang, Sichuan 621010, China

⁴Department of Pathology, 404 Hospital, Mianyang, Sichuan 621010, China

Correspondence should be addressed to Yongxiang Gao; 1124318366@qq.com

Received 9 June 2022; Revised 19 July 2022; Accepted 28 August 2022; Published 12 September 2022

Academic Editor: Kuruva Lakshmana

Copyright © 2022 Mingwei Qin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Ubiquitous computing and artificial intelligence are widely used in the field of Wise Information Technology of Med. It promotes the development of intelligent diagnosis and treatment, as well as reducing the workload of doctors' diagnosis and promoting the improvement of the diagnosis level of medical institutions in remote areas. The accurate detection of cervical cytopathology is related to the precise treatment and rehabilitation of patients. However, the low rate of accuracy in the existing cervical cytopathology image detection cannot satisfy the application needs of intelligent diagnosis. In this paper, a dual path network efficient detection method for cervical cytopathology (the proposed DSRNet50) is proposed, which is based on the deep learning method, and combines residual structure and dense connection. The proposed DSRNet50 is mainly based on residual structure, supplemented by dense connection paths, which improves the utilization of features, and maintains the ability of exploring new features by reducing feature redundancy. Meanwhile, the proposed DSRNet50 leverages packet convolution to reduce the computation burden of the network and eliminate the over fitting phenomenon of the network. The proposed DSRNet50 further exploits channel attention mechanism to recalibrate important features to suppress the propagation of irrelevant information. We use the Herlev dataset to verify the proposed DSRNet50 in the terms of the detection accuracy, parameter quantity, and computing complexity. The experiment results are presented to show the achievable performance of the proposed DSRNet50.

1. Introduction

In recent years, ubiquitous computing and artificial intelligence technology have developed rapidly and have been widely used in the medical field [1], the automobile field [2], and the UAV communication field [3]. In the medical field, with the development of Wise Information Technology of Med (WITMED), people can obtain digital medical services anytime and anywhere in an environment full of computing and communication capabilities. The intelligent diagnosis method based on medical knowledge, combined with computer technology, image processing technology, and artificial intelligence, as the key technology in WITMED, can automatically identify suspicious lesions, improve doctors' work efficiency, reduce the misdiagnosis rate and the

missed diagnosis rate, improve the diagnosis and treatment level, and solve the problem of lack of medical resources in remote areas. As the most common and effective means of early diagnosis and prevention of cervical cancer and its precancerous lesions [4], cervical cell screening plays an important role in the diagnosis and prevention of cervical cell lesions. Before the emergence of intelligent diagnosis systems, the main method used was manual film reading. Manual film reading has the disadvantages of long diagnosis time, low efficiency [5], and high dependence of accuracy on the level of professional technicians. The sensitivity of film readers is only about 65% due to fatigue, skill level, and subjective interpretation [6], which is difficult to meet the needs of cervical cancer screening in the current society. With the emergence of intelligent diagnosis technology, especially in

cell classification, intelligent diagnosis technology has become a research hotspot of WITMED. Because it adds the experience and knowledge of highly skilled doctors, and it has the characteristics of high film reading efficiency and not affected by working hours, doctors in remote areas can also rely on the intelligent diagnosis system to treat difficult and miscellaneous diseases. Therefore, intelligent diagnosis technology has become a research hotspot of WITMED.

In order to meet the requirements of intelligent diagnosis, it is necessary to further improve the accuracy of cervical cytopathology detection. At the same time, in order to meet the low-cost characteristics of ubiquitous computing, it is necessary to reduce the increased computing resources as much as possible. Based on the above requirements, this paper proposed DSRNet50 based on dual path network, which takes the residual structure as the main body, adds grouping convolution to reduce the amount of parameters, and then uses the channel attention mechanism to help the network learn important feature information, suppress unimportant feature information, realize feature recalibration, and use dense connection to build a dual path structure, so that the network can continue to explore new features. It not only increases very few computing resources but also greatly increases the accuracy and meets the needs of intelligent diagnosis and ubiquitous computing.

The sections of this paper are mainly arranged as follows:

The first section introduces the development and advantages of pervasive computing and artificial intelligence technology in the medical field, especially the importance of cervical cell reading in cervical cancer diagnosis, illustrates the limitations of traditional reading methods with examples, and then introduces the advantages of intelligent diagnosis and treatment technology, which has the problem of insufficient accuracy in cervical cancer cell reading. Therefore, the DSRNet50 network proposed in this paper is introduced. It can effectively improve the accuracy of film reading and introduces the contribution of this work

The second section introduces the related work of cervical cytopathology recognition and compares it with the work of this paper

The third section introduces the DSRNet50 proposed in this paper. The principle and structure of DSRNet50 are introduced in detail

The fourth section introduces the experimental and simulation results, including dataset selection, dataset preprocessing, evaluation parameter analysis, and simulation results. The results show that the network proposed in this paper has better recognition effect of cervical cytopathology than other networks and only needs to increase a small amount of calculation

The fifth section summarizes and prospects the article

2. Related Works

As an algorithm applied in intelligent diagnosis technology, cervical cell image classification algorithms can assist doctors to complete cervical cell screening. In the early cervical cell image classification algorithm, the morphology and texture

of cells were mainly selected manually, such as cell shapes, areas, dyeing depths, the roundness of nucleus, and the nucleocytoplasmic ratio [7], and then a specific classifier or a neural network was used for automatic classification. However, this method is limited by the labeling accuracy and the segmentation accuracy of cell image features, and its accuracy fluctuates greatly; so, it is not enough to be included in the primary screening method [8]. Although the cell segmentation algorithm has been optimized [9–13], the similarity between cell clusters and normal/abnormal cells still hinders the accurate segmentation of cells. Once the segmentation deviates [11], the classification accuracy will decline. For example, Lassouaoui and Hamami. used the region growth algorithm for cervical cell image segmentation [14]. Plissiti and Nikou extracted the features of the nucleus, used the principal component analysis algorithm to reduce the dimension, and then used the fuzzy mean algorithm as the classifier [15]. Genctav et al. and Nayar and Wilbur grouped and sorted according to TBS (cervical cell standard classification system) [16, 17] and used the hierarchical clustering algorithm to construct decision trees for classification. Zhao et al. used a three-stage strategy to classify cervical cancerous cells [18], extracted 120 Witt's sign, and used it for the training of linear classifiers.

As a deep learning model, the convolution neural network convolutes the input data, reduces the complexity of data, automatically obtains image features, avoids manual design features, and implicitly extracts image features, which just makes up for the low accuracy caused by large data deviation. In 2012, Krizhevsky et al. applied the deep convolution neural network model AlexNet to ImageNet dataset [19], which improved its accuracy by about 10%. Zhang et al. applied convolutional neural network to the classification task of normal and abnormal cervical cells for the first time [20] and migrated the model in the ImageNet large-scale visual recognition challenge to the classification model of cervical cell image but did not make detailed classification according to the lesion level. Since then, many studies on cervical cell classification have emerged, such as Xu et al.'s image of cervical dysplasia combined with multimodal data and deep learning for cancer diagnosis [21]. Sarwar et al. established a standard database based on the cell images and cervical cytopathological knowledge and used convolutional neural network to realize the digitization of screening [22]. The above methods can realize cancer recognition. In addition, Tareef and others have also achieved good results in the recognition of overlapping cell images by using dynamic shape modeling method combined with deep learning [23]. Although the above methods have improved the accuracy to a certain extent, there are still deficiencies in sensitivity and recall. To be applied to intelligent diagnosis, the indicators need to be further improved.

The network proposed in this paper focuses on the defects of low sensitivity and recall rate in the above research and takes into account the problem of computing resources for actual deployment. The follow-up experimental results show that the indicators are significantly improved than the previous literature and can effectively identify cervical cytopathology.

3. The Proposed DSRNet50

The proposed DSRNet50 introduced residual network, block convolution, dense connection path, and channel attention mechanism. This section will first introduce the principle of each part of the proposed DSRNet50, briefly describe the reason why it is added to the proposed DSRNet50, and then introduce the overall network structure of the proposed DSRNet50.

3.1. Principle of DSRNet50

3.1.1. The First Component of DSRNet50: ResNet50. For deep neural networks, it is generally believed that the deeper the number of network layers, the stronger the nonlinear expression ability of the network, and the network can learn more features; so, the accuracy should be higher. However, it is found that this is not the case in the experimental process. With the increase of the number of layers, the training effect of the traditional convolutional neural network will become worse, and the accuracy will become lower. In order to solve this problem, in 2015, He and others proposed the deep residual network ResNet50 [24]. ResNet50 is a deep residual network which has 50 network layers.

The design idea of ResNet50 is as follows: assuming that a network A can work well, and now there is a network B which is deeper than A , let the former part of B be exactly the same as A , so that the latter part only needs to realize an identity mapping, so that B can at least obtain the same performance as A . As shown in Figure 1, assuming that the input of A certain neural network is x and the expected output is $H(x)$, i.e., $H(x)$ is the expected complex potential mapping, and it will be more difficult to learn such a model.

Back to the previous hypothesis, when the learning accuracy is saturated, or when the error of the lower layer is found to be larger, the next learning goal can be transformed into the learning of identity mapping, even if the input x is similar to the output $H(x)$, so that the accuracy of the network will not decline in the later level.

In the residual network structure diagram shown in Figure 1, the input x is directly transmitted to the output as the initial result through “shortcut connections,” and the output is $H(x) = F(x) + x$. if $F(x) = 0$, then $H(x) = x$, which becomes the identity map mentioned above. At this time, the network is equivalent to changing the learning goal. It is no longer learning a complete output, but the difference between the learning goal value $H(x)$ and x , i.e., the so-called residual value $F(x)$. Therefore, the subsequent training goal is to approximate the residual value to 0, so that the accuracy will not decrease with the deepening of the network.

Although the deep residual network solves the problem that it is difficult to train the deep-seated network and improves the performance of the deep neural network to a certain extent, its accuracy still cannot meet the application in the medical field.

3.1.2. The Second Component of DSRNet50: Block Convolution. Block convolution is to group different characteristic images of the input layer and then use different con-

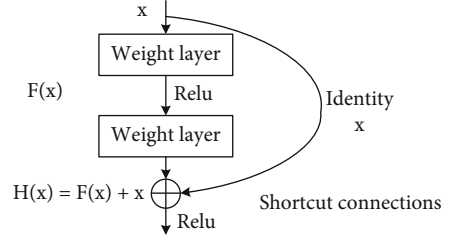


FIGURE 1: Structural diagram of residual block.

volution to check each group for convolution, which can reduce the amount of convolution calculation, reduce the correlation between channels, and prevent over fitting. Because ordinary convolution is convolution on all input characteristic graphs, i.e., full channel convolution, which is a channel dense connection, packet convolution is a channel sparse connection.

Figure 2 represents the standard convolution operation. If the dimension of the input characteristic drawing is $H \times W \times c_1$, the size of convolution kernel is $h_1 \times w_1 \times c_1$, the size of the output characteristic diagram is $H \times W \times c_2$, and the parameter quantity of standard convolution layer is $h_1 \times w_1 \times c_1 \times c_2$.

Figure 3 represents a packet convolution operation. If the input feature map is divided into g groups according to the number of channels, the size of each group of input feature map is $H \times W \times c_1/g$. The corresponding convolution kernel size is $h_1 \times w_1 \times c_1/g$. The size of each group of output characteristic diagram is $H \times W \times c_2/g$. Splice the results of group g to obtain the final size of $H \times W \times c_2$ output characteristic diagram. The parameter quantity of the block convolution layer is $h_1 \times w_1 \times c_1/g \times c_2/g \times g = h_1 \times w_1 \times c_1 \times c_2 \times 1/g$.

On the characteristic graph of ordinary convolution output, each point is represented by the input characteristic graph $h_1 \times w_1 \times c_1$ points that are calculated, and on the characteristic graph of the block convolution output, each point is calculated by the input characteristic graph of $h_1 \times w_1 \times c_1/g$ points. Therefore, the parameter quantity of ordinary convolution is g times of block convolutions.

However, considering the lack of information exchange between channels after adding block convolution, which may lead to the loss of information, 1×1 convolution kernel and channel attention mechanism are adopted after channel splicing to ensure the information exchange between the characteristic maps of different groups, as is shown in Figure 4.

3.1.3. The Third Component of DSRNet50: Channel Attention Mechanism. The channel attention mechanism helps the network locate the channels that have extracted important features and give higher weights. In this paper, the Se module [25] is used to recalibrate the features of each channel, i.e., the redistribution of weights, to help the network learn important feature information, as is shown in Figure 5.

(1) *Squeeze.* Feature compression is carried out in the spatial dimension to turn each two-dimensional feature channel

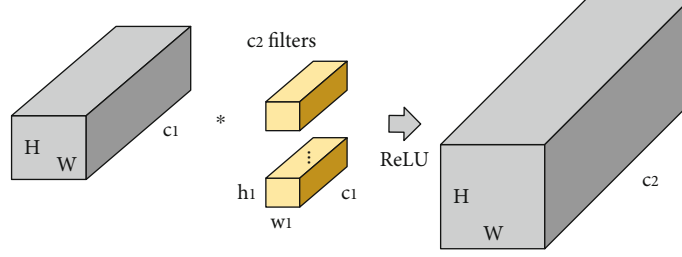


FIGURE 2: Ordinary convolution process.

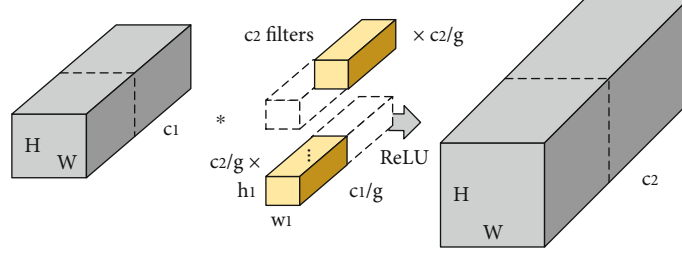


FIGURE 3: Grouped convolution process.

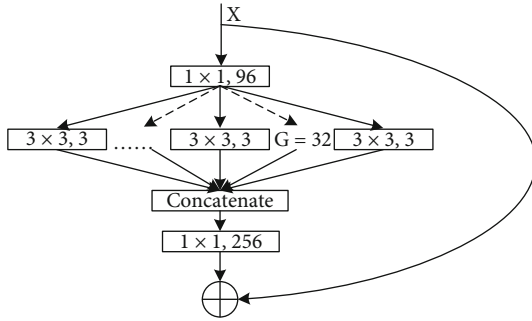


FIGURE 4: Structural diagram of grouped convolution.

into a real number, which has a global receptive field to some extent, and the output dimension matches the number of input feature channels. It represents the characteristic channel responding to the global distribution and enables the layer close to the input to obtain the global receptive field.

This operation uses the global average pooling, i.e., the sum of all pixels of a feature map is averaged to obtain a value, which is used to represent a feature map, and the input characteristics of $H \times W \times C$ are synthesized as the feature description of $1 \times 1 \times C$. The calculation of a feature map is as follows:

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c. \quad (1)$$

(2) *Excitation*. After squeeze operation, the network only gets a global description, and this result cannot be used as the weight of the channel. The exception operation can comprehensively obtain the dependency relationship between channels. This operation includes two full connection layers and two activation functions; so, it has more nonlinearity to

fit the complex correlation between channels. It can integrate all the input characteristic information well and map the input to the 0~1 interval to realize the gating mechanism. The mathematical expression of exception operation is shown in equation (2):

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)). \quad (2)$$

In the formula, z is the global description obtained by the squeeze operation, δ is the ReLU function, W_1 and W_2 are two full connection layers, in which $W_1 \in R^{C/r \times C}$, $W_2 \in R^{C/r \times C}$, and r are the scaling parameters, which is mainly used to reduce the computation complexity and parameter quantity of the network, and also affect the experimental accuracy. In this paper, r is selected as 16, taking into account the accuracy and parameter quantity.

(3) *Reweight*. The reweight operation takes the result of the exception output as the importance of each feature channel after feature selection, i.e., the weight coefficient, and then multiplies it to each feature channel one by one to complete the recalibration of the original feature in the channel dimension. The calculation formula of this operation is as follows:

$$\tilde{X} = F_{scale}(u_c, S_C) = S_C * u_c. \quad (3)$$

The characteristic graph u_c of each channel is multiplied by the corresponding scalar S_C . The element S_C in the vector obtained in the exception operation is used as the weight coefficient of the corresponding channel.

3.1.4. *The Fourth Component of DSRNet50: Dense Connection Path*. Inspired by DenseNet [26], this paper will

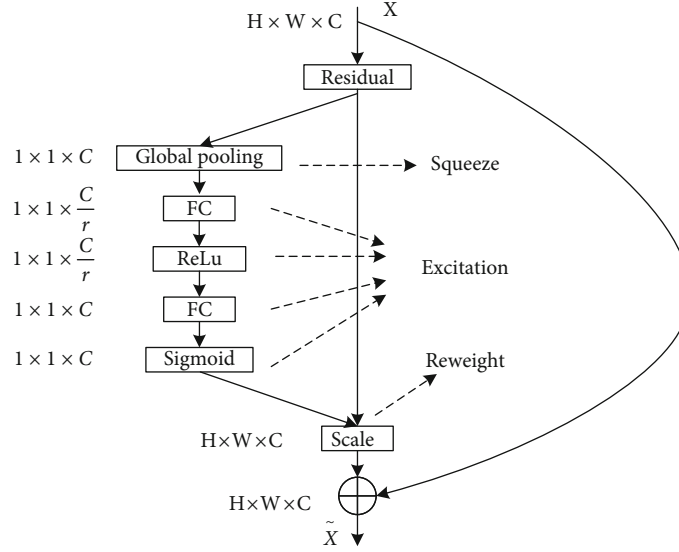


FIGURE 5: Structure diagram of SE module.

adopt a new dense connection mode, which is similar to the feature fusion of local two paths, combined with the residual network structure to reuse the features, so as to learn and create new features.

(1) *DenseNet and Higher Order Recurrent Neural Network.* HORNN (higher order recurrent neural network) can be generalized to formula (4). In the formula, h^k is the current state, h^t represents the t -th state in the structure ($t < k$), $f_t^k(\cdot)$ represents the process of extracting features related to the k -th state from h^t , x^t represents the input data of the t -th state, where $x_0 = h_0$, and $g^k(\cdot)$ represents that the current state h^k is obtained by aggregating all previous states and then processing them with $g^k(\cdot)$. If $f_t^k(\cdot) = f_t(\cdot)$ and $g^k(\cdot) = g(\cdot)$ for any k and t , formula (4) is an ordinary HORNN.

$$h^k = g^k \left[\sum_{t=0}^{k-1} f_t^k(h^t) \right]. \quad (4)$$

The output of the previous layers is spliced in the feature dimension, and then 1×1 convolution is performed. This operation can be equivalent to doing 1×1 convolution on the direct line of each layer first. The convolution coefficients of 1×1 are different on each straight line and then were arithmetically added, which fully corresponds to formula (4). Therefore, DenseNet is a special HORNN when $f_t^k(\cdot) = f_t(\cdot)$ and $g^k(\cdot) = g(\cdot)$ are not satisfied. Among them, f_{k-1}^k , f_{k-2}^k , f_{k-3}^k , ... are different, which means that the features extracted from the front layer are no longer simply reused by the back layer, but create new features. However, the features extracted by convolution in the back layer of this structure are likely to have been extracted by the front layer.

(2) *DenseNet and Residual Network.* Deform formula (4), introduce an intermediate variable r^k , and then assume that

for any t and k , $f_t^k(\cdot) = f_t(\cdot)$, formula (4) can be changed into the expression of formulas (5) and (6):

$$r^k \triangleq \sum_{t=1}^{k-1} f_t(h^t) = r^{k-1} + f_{k-1}(h^{k-1}), \quad (5)$$

$$h^k = g^k(r^k). \quad (6)$$

Combining the above two expressions together gives formula (7), where $\Phi_{k-1}(r_{k-1}) = f_{k-1}(g_{k-1}(\cdot))$. It can be seen that this is a residual expression:

$$\begin{aligned} r^k &= r^{k-1} + f_{k-1}(h^{k-1}) = r^{k-1} + f_{k-1}(g^{k-1}(r^{k-1})) \\ &= r^{k-1} + \phi^{k-1}(r^{k-1}). \end{aligned} \quad (7)$$

Because formula (5) is derived from formula (4) on the premise of $f_t^k(\cdot) = f_t(\cdot)$, ResNet can be considered as a special expression of DenseNet when $f_t^k(\cdot) = f_t(\cdot)$.

3.2. *The Proposed DSRNet50 Structure.* In this paper, the local dual path architecture is adopted, and the residual structure is still used as the main network, which is composed of stacking multiple DSR residual blocks. As shown in Figure 6, first pass through a convolution layer of 1×1 , connecting a block convolution layer of 3×3 , and then splice the output of the block convolution. The spliced output will be divided into two channels, and one through the convolution layer of 1×1 is sent to the SE module for feature recalibration and added to the residual path in the form of elements. Another way through another convolution kernel of 1×1 is connected with the dense connection path.

The residual network reuses the features extracted from the previous layer. In addition to these directly connected reuse features, the features really extracted by convolution

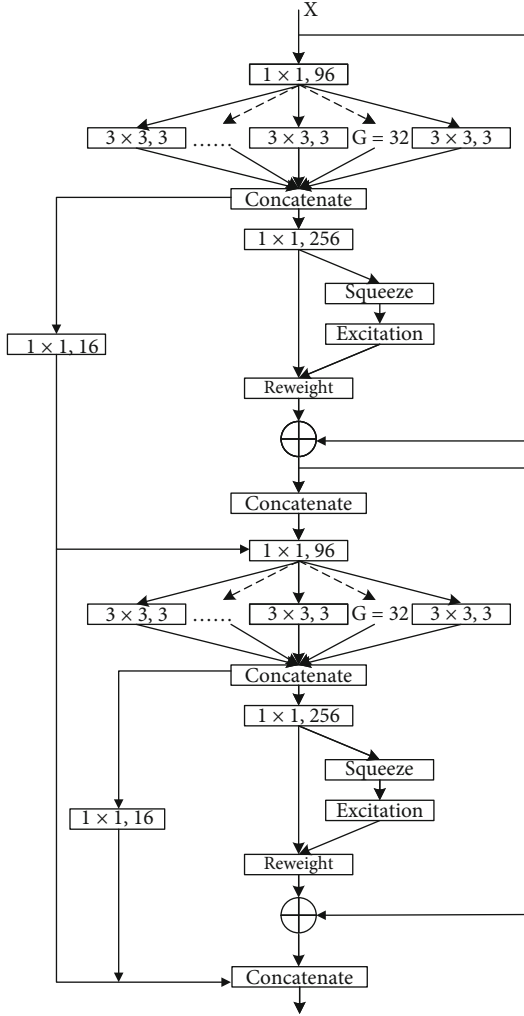


FIGURE 6: Structural diagram of DSR residual block.

are basically new features that have not been extracted before, and the redundancy of the extracted features is relatively low. The features extracted through dense connection are no longer simply reused by the later layer but create new features. However, this structure will lead to that the features extracted by the later convolution layer are likely to have been extracted by the previous layer, and the extracted features have high redundancy. DSRNet combined the advantages of the two networks and added packet convolution and channel attention mechanism, so that it can have good computational characteristics while maintaining high reuse rate and low redundancy of features and continuously explore new features and suppress useless features.

4. Experiment and Analysis

This section mainly describes the datasets cited in this paper, as well as the preprocessing methods for datasets to prevent over fitting. Finally, taking the general evaluation method as the index, the advantages and disadvantages of different algorithms in the index are tested.

TABLE 1: Herlev cervical cell dataset.

Type	Name	Number of images
Normal	Normal superficile	74
	Normal intermediate	70
	Normal columnar	98
Abnormal	Lightweight dyspastic	182
	Moderate dayspatial	146
	Server dayspatial	197
	Carcinoma	150

4.1. Dataset. Herlev cervical cell dataset [11] was used by Herlev University Hospital (HUH) and Technical University of Denmark (TUD) with a resolution of $0.201 \mu\text{M}$ digital camera and microscope to collect the new version of Herlev Pap smear dataset. There are 917 single-cell images in the dataset, including 242 normal cells and 675 abnormal cells. The specific number is shown in Table 1.

It can be seen from Figure 7 that the nuclear cytoplasmic ratio and nuclear cytoplasmic brightness are obvious characteristic differences. The images of seven types of cervical cells are different in shape, among which Figures 7(a)–7(c) are normal cells and Figures 7(d)–7(g) are diseased cells. Most normal cells are quite different from abnormal cells, columnar epithelial cells, squamous cell carcinoma in situ, and severely squamous intraepithelial diseased cells, and the similarity between middle squamous cells and epithelial squamous cells is high. It is difficult to distinguish.

4.2. Dataset Preprocessing. Data expansion is a common method in deep learning, that is, making a certain transformation on the image data to increase the number of input data, which can prevent the overfitting phenomenon due to less data. However, due to the particularity of cervical cells, the common color transformation, scale transformation, and noise transformation in data expansion will affect the original characteristics of cervical cells; so, it is not applied to the expansion of this type of dataset. All images are processed in strict accordance with the standard of TBS2014. In this paper, the methods of flip and rotation are used to expand, and then all the images in the dataset are processed into 224×224 pixel images. The number of cell images in the expanded dataset is shown in Table 2.

4.3. Evaluating Indicator. This paper takes the expanded dataset as the training set and the original dataset as the test set. The evaluation method is generally to compare the real situation of patients with the diagnosis of the system. The diagnosis results include four categories, namely,

True negative (T_N): the true condition is no lesion, and the diagnosis is no lesion

False negative (F_N): the real situation is lesion, and the diagnosis is no lesion

False positive (F_p): the real situation is no lesion, and the diagnosis is lesion

True positive (T_p): the real situation is lesion, and the diagnosis is lesion

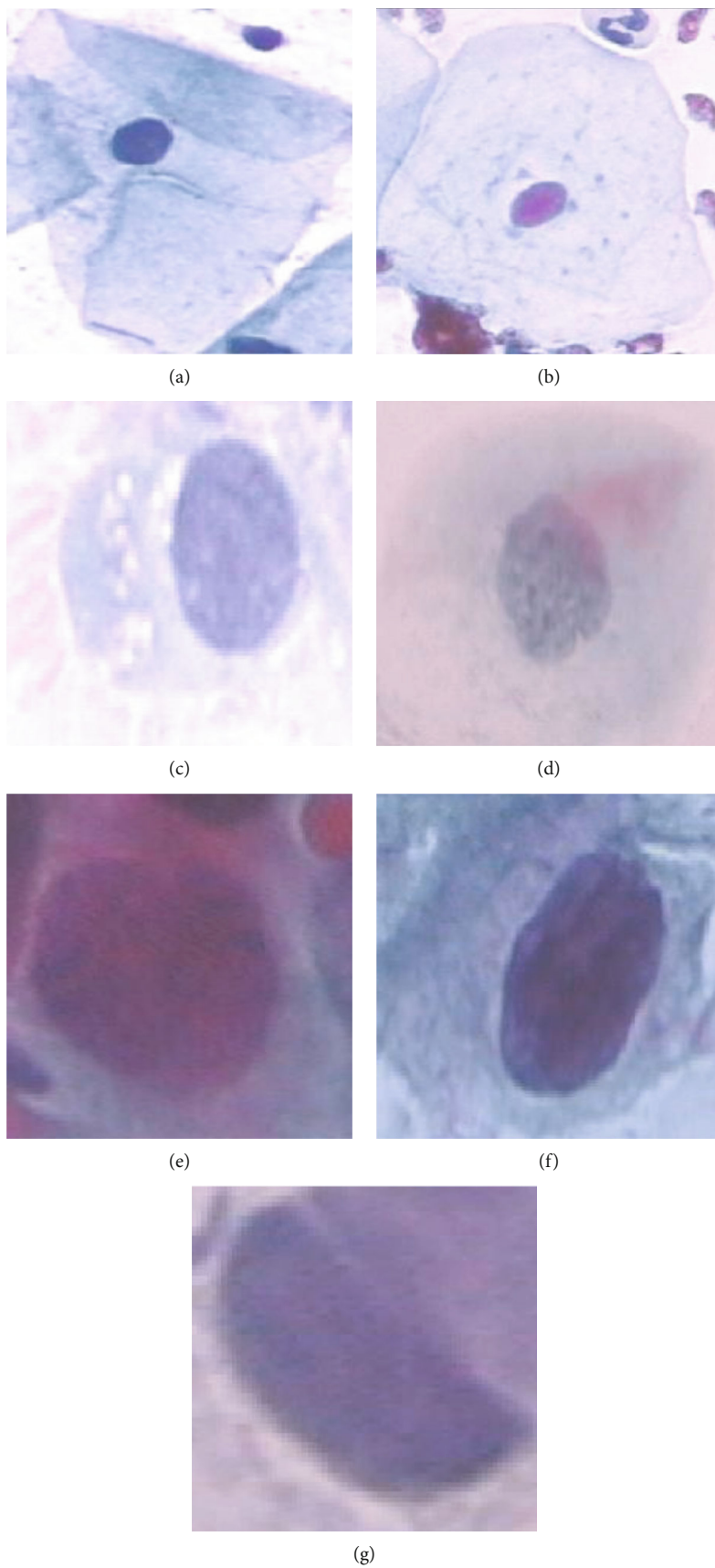


FIGURE 7: 7 types of cervical cell images: (a) normal superficile, (b) normal intermediate, (c) normal columnnar, (d) lightweight dyspastic, (e) moderate dayspatial, (f) server dayspatial, and (g) carcinoma.

TABLE 2: Extended Herlev cervical cell dataset.

Type	Name	Number of images
Normal	Normal superficile	370
	Normal intermediate	350
	Normal columnar	490
Abnormal	Lightweight dyspastic	910
	Moderate dayspatial	730
	Server dayspatial	985
	Carcinoma	750

TABLE 3: Standard formula of evaluation index.

Evaluation index	Formula
ACC	$T_p + T_N / T_p + T_N + F_p + F_N$
P	$T_p / T_p + F_p$
R	$T_p / T_p + F_N$
F1	$2 \times P \times R / P + R$

The evaluation indicators also include four categories: accuracy (ACC), that is, the proportion of correctly classified samples in the total samples; the accuracy rate (P) is how many samples predicted as positive examples are real positive examples; recall rate (R), that is, how many positive examples in the sample are predicted correctly, i.e., sensitivity; and F1-score is the weighted harmonic average value of P and R , and it is an index proposed considering the unbalanced number of samples in each classification. When F1-score is high, the expression method is more effective. Since this paper realizes seven classifications of cervical cell images, the total index of the model needs to be obtained after the sum of the indexes of each classification. The calculation formula of four indicators is shown in Table 3:

4.4. Experimental Results and Analysis. All experimental results in this section are obtained in the same environment. Experimental environment configuration is as follows: the operating system is windows 8.1 \times 64. The CPU is Intel Xeon 8-core 2620v4 \times 2; GPU is NVIDIA GTX 1080ti and 16 GB memory. The framework is TensorFlow 3.0. The programming language is Python 3.6.

Test seven classification performances of Herlev dataset on the proposed DSRNet50. Set epoch to 100, which means 100 rounds of iteration. The scaling parameter is $R = 16$, and the grouping convolution parameter is $G = 32$. As shown in Figure 8, the proposed DSRNet50 gradually converges with the increase of the number of iterations. After about 20 full set iterations, the curve tends to be flat.

In order to fully illustrate the effectiveness of this method, the structure of different methods will be adjusted based on the original RESNet50 network, and a comparative experiment will be carried out on the Herlev dataset. It can be seen from Table 4 that the results of several improved networks using different methods are compared. GC represents packet convolution, Se represents channel attention module, and D represents dense connection. N1 is the orig-

inal RESNet50 network, N2 to N4 are improved networks using different methods, and N5 is the proposed DSRNet50 network in this paper. It can be seen from the table that after adding different methods, the performance of the network has been improved to varying degrees. Compared with N1, N2 added with channel attention module has improved the ACC, P , and R by 2.9%, 2.1%, and 2.5%, respectively, and then N3 has further improved by 1.2%, 1.3%, and 0.6% after adding grouping convolution on the basis of N2. After adding dense connection to N2, N4 increased by 0.6%, 0.5%, and 0.7% compared with N3. Compared with N1, the ACC, P , and R of algorithm N5 in this paper are improved by 6.6%, 5.5%, and 5.2%, respectively, which shows the effectiveness of this method in the detection of cervical lesion cells.

In terms of parameters and calculation amount, since the number of channels in the last layer is 2048, the model has a large increase in parameters; so, the SE module is removed in the last layer, which has little impact on the ACC. As shown in Table 4, N5 has an increase of about 5% in parameters compared with N1, and the amount of calculation has increased by about 7%. It can be seen that the proposed DSRNet50 improves the accuracy and only adds small parameters and amount of calculation, which is acceptable.

Table 5 shows the detailed indicators of seven classifications of the proposed DSRNet50 in the test set. It can be seen from the table that the proposed DSRNet50 shows good classification performance and can well distinguish various types of pathological cells, but the classification effect of severe squamous intraepithelial lesions and squamous cell carcinoma in situ is slightly poor, indicating that the characteristics of these two types of cells are highly similar to those of other cells, so that the network cannot distinguish their pathological types sometimes.

In order to further illustrate the performance of this algorithm in the field of cervical cytopathy detection, this algorithm is compared with the performance of other references on the Herlev dataset. Because most algorithms on the Herlev dataset are the two classifications of cervical cells with or without lesions, they cannot be directly compared in performance, but for cells, multiclassification is more persuasive than two classifications. It can better illustrate the accuracy of the algorithm. The comparison results are shown in Table 6. The algorithm in this paper shows comparable or even better performance with other two classification and seven classification algorithms, which is enough to illustrate the effectiveness of this method.

To sum up, for this paper, the difficulty is to achieve excellent classification effect in multiclassification of the same thing. In medical assisted diagnosis, accuracy and recall are the most critical evaluation indicators. According to the experimental results in this section, this method has better classification performance than other two classification algorithms in the seven classification tasks of cervical cells. Literature [13] uses genetic algorithm and 1-neighbor algorithm for classification. Literature [11] extracts feature classification after presegmenting images. Literature [27] uses artificial neural network ANN for classification. Literature [11] designed and used three classifiers to vote for

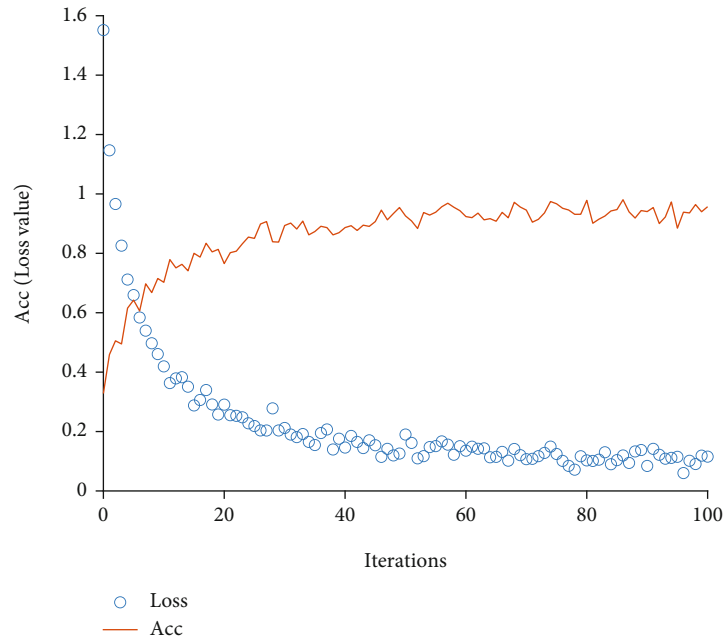


FIGURE 8: Accuracy/loss value curve.

TABLE 4: Model evaluation indicators of different methods.

	Method		Evaluation indicators						Model size	
	ResNet50	GC	SE	D	ACC/%	P /%	R /%	F1/%	Parameters	FLOPs
N1	√				92.3	93.7	93.6	93.5	23.5×10^6	3.86×10^9
N2	√		√		95.2	95.8	96.1	95.8	24.1×10^6	3.87×10^9
N3	√	√	√		96.4	97.1	96.7	96.9	23.7×10^6	3.97×10^9
N4	√		√	√	97.0	97.6	97.4	97.5	25.1×10^6	4.12×10^9
N5	√	√	√	√	98.9	99.2	98.8	99.0	24.7×10^6	4.14×10^9

TABLE 5: The proposed DSRNet50 test set classification evaluation indicators.

Cell type	P /%	R /%	F1/%
Normal superficile	100.0	97.7	99.3
Normal intermediate	99.3	100.0	99.3
Normal columnar	100.0	99.0	99.5
Lightweight dyspastic	99.5	98.8	98.9
Server dayspatial	99.4	98.6	99.2
Server dayspatial	97.0	99.5	97.8
Carcinoma	97.2	98.7	97.7

classification. Literature [20] uses convolutional neural network to automatically extract features for classification. Although literature [11, 13, 27] has high accuracy, it needs to design features manually. The accuracy of literature [11, 27, 28] is low, which shows that the ability to correctly predict positive cases is poor. The average error rate of this algorithm on Herlev dataset is 1.5%. Compared with other binary classification algorithms, this algorithm has more practical value in the field of auxiliary diagnosis. In terms of seven categories, the proposed algorithm improves ACC,

TABLE 6: Performance comparison between the proposed DSRNet50 and other algorithms.

Method	Number of categories	ACC/%	P /%	R /%
Marinakis et al. [13]	2	96.8	92.1	98.5
Jantzen et al. [11]	2	93.6	79.3	98.8
Bora et al. [27]	2	96.5	89.7	99.3
Kim et al. [28]	2	93.6	81.3	98.3
Zhang et al. [20]	7	98.3	98.3	98.2
Ours	7	98.9	99.2	98.8

P , and R by 0.5%, 0.9%, and 0.5%, respectively. For tens of thousands of cases in the future, a considerable number of families can be saved from later treatment [20]. In general, the effectiveness of the proposed DSRNet50 is proved by the experiment.

5. Conclusion

Aiming at the problems of low efficiency and low accuracy in the existing detection methods of cervical cytopathy, this paper proposes an improved deep convolution neural

network DSRNet, which adopts a dual path structure, takes the residual path as the main body, and adds local dense connection paths, so that the network can constantly explore new features while maintaining the high reuse rate and low redundancy of features. In addition, packet convolution is added to the main network to reduce the amount of calculation and better prevent the phenomenon of overfitting. Then, the SE module is used for feature recalibration to make better use of important features and suppress useless features. Experimental results show that this algorithm can detect seven kinds of normal and diseased cells with high accuracy on Herlev dataset. Compared with other algorithms, it has good performance improvement and fills the defects of ubiquitous computing in WITMED and intelligent diagnosis. In this follow-up work, we can collect more data, improve the reliability of experimental results, consider hardware acceleration research according to the requirements of practical application, and design a practical detection algorithm that takes into account both accuracy and computational complexity.

Data Availability

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

Special Project 3 for Major Instrument and Equipment Development of Key R&D Plan (2016YFF0104003). We thank Mingliang Gou, Yuancheng Yao, and Ning Zhou for the completion of the experiment and Yongxiang Gao for the completion of this article.

References

- [1] A. Jimenez-Molina, J. Gaete-Villegas, and J. Fuentes, "Pro-FUSO: business process and ontology-based framework to develop ubiquitous computing support systems for chronic patients' management," *Journal of Biomedical Informatics*, vol. 82, pp. 106–127, 2018.
- [2] E. Bran, E. Bautu, D. F. Sburlan, C. M. Puchianu, and D. M. Popovici, "Ubiquitous computing: driving in the intelligent environment," *Mathematics*, vol. 9, no. 21, p. 2649, 2021.
- [3] L. Chen, R. Zhao, K. He, Z. Zhao, and L. Fan, "Intelligent ubiquitous computing for future UAV-enabled MEC network systems," *Cluster Computing*, vol. 25, no. 4, pp. 2417–2427, 2022.
- [4] M. L. Hutchinson, D. J. Zahniser, M. E. Sherman et al., "Utility of liquid-based cytology for cervical carcinoma screening," *Cancer*, vol. 87, no. 2, pp. 48–55, 1999.
- [5] A. S. Rustagi, A. Kamineni, S. Weinmann, S. D. Reed, P. Newcomb, and N. S. Weiss, "Cervical screening and cervical cancer death among older women: a population-based, case-control study," *American Journal of Epidemiology*, vol. 179, no. 9, pp. 1107–1114, 2014.
- [6] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2018," *CA: a Cancer Journal for Clinicians*, vol. 68, no. 1, pp. 7–30, 2018.
- [7] R. Nayar and D. C. Wilbur, *The Bethesda System for Reporting Cervical Cytology : Definitions, Criteria, and Explanatory Notes*, Springer, Cham, 2015.
- [8] H. C. Kitchener, R. Blanks, G. Dunn et al., "Automation-versus manual reading of cervical cytology (MAVARIC): a randomised controlled trial," *The Lancet Oncology*, vol. 12, no. 1, pp. 56–64, 2011.
- [9] L. Kuan, *Research on Segmentation, Texture Extraction and Classification Methods for Cell Images*, ph.D. dissertation, National University of Defense Technology, China, 2012.
- [10] E. Martin, *Pap-Smear Classification*, ph.D. dissertation, Technical University of Denmark, Denmark, 2003.
- [11] J. Jantzen, J. Norup, G. Dounias, and B. Bjerregaard, "Pap-smear benchmark data for pattern classification," in *Proc.NI-SIS*, pp. 1–9, Tenerife, Spain, 2005.
- [12] T. Chankong, N. Theera-Umpon, and S. Auephanwiriyakul, "Automatic cervical cell segmentation and classification in pap smears," *Computer Methods and Programs in Biomedicine*, vol. 113, no. 2, pp. 539–556, 2014.
- [13] Y. Marinakis, G. Dounias, and J. Jantzen, "Pap smear diagnosis using a hybrid intelligent scheme focusing on genetic algorithm based feature selection and nearest neighbor classification," *Computers in Biology and Medicine*, vol. 39, no. 1, pp. 69–78, 2009.
- [14] N. Lassouaoui and L. Hamami, "Genetic algorithms and multi-fractal segmentation of cervical cell images," in *Proceeding.Seventh International Symposium on Signal Processing and Its Applications*, vol. 2, pp. 1–4, Paris, France, July 2003.
- [15] M. E. Plissiti and C. Nikou, "Cervical cell classification based exclusively on nucleus features," in *International Conference Image Analysis and Recognition*, pp. 483–490, Springer, Berlin, Heidelberg, 2012.
- [16] A. Gençtav, S. Aksoy, and S. Önder, "Unsupervised segmentation and classification of cervical cell images," *Pattern Recognition*, vol. 45, no. 12, pp. 4151–4168, 2012.
- [17] R. Nayar and D. C. Wilbur, "The pap test and Bethesda 2014," *Cancer Cytopathology*, vol. 123, no. 5, pp. 271–281, 2015.
- [18] L. Zhao, K. Li, J. Yin, Q. Liu, and S. Wang, "Complete three-phase detection framework for identifying abnormal cervical cells," *IET Image Processing*, vol. 11, no. 4, pp. 258–265, 2017.
- [19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [20] L. Zhang, le Lu, I. Noguees, R. M. Summers, S. Liu, and J. Yao, "DeepPap: deep convolutional networks for cervical cell classification," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 6, pp. 1633–1643, 2017.
- [21] T. Xu, H. Zhang, X. Huang, S. Zhang, and D. N. Metaxas, "Multimodal deep learning for cervical dysplasia diagnosis," in *International conference on medical image computing and computer-assisted intervention*, Springer, Cham, 2016.
- [22] A. Sarwar, J. Suri, V. Sharma, and M. Ali, "Novel benchmark database of digitized and calibrated cervical cells for artificial intelligence based screening of cervical cancer," *Gynecology & Obstetrics*, vol. s3, 2017.

- [23] A. Tareef, Y. Song, H. Huang et al., "Optimizing the cervix cytological examination based on deep learning and dynamic shape modeling," *Neurocomputing*, vol. 248, pp. 28–40, 2017.
- [24] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Las Vegas, NV, USA, June 2016.
- [25] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, June 2018.
- [26] Y. Chen, J. Li, H. Xiao, X. Jin, S. Yan, and J. Feng, "Dual path networks," *Advances in Neural Information Processing Systems*, vol. 30, 2017.
- [27] K. Bora, M. Chowdhury, L. B. Mahanta, M. K. Kundu, and A. K. das, "Automated classification of pap smear images to detect cervical dysplasia," *Computer Methods and Programs in Biomedicine*, vol. 138, pp. 31–47, 2017.
- [28] K.-B. Kim, D. H. Song, and Y. W. Woo, "Nucleus segmentation and recognition of uterine cervical pap-smears," in *International Workshop on Rough Sets, Fuzzy Sets, Data Mining, and Granular-Soft Computing*, pp. 153–160, Springer, Berlin, Heidelberg, 2007.