



HAL
open science

A HEP-2 Cell Image Classification Model Based on Deep Residual Shrinkage Network Combined with Dilated Convolution

Chen Wang, Tao He, Jiansheng Liu, Dapeng Li, Yingyou Wen

► **To cite this version:**

Chen Wang, Tao He, Jiansheng Liu, Dapeng Li, Yingyou Wen. A HEP-2 Cell Image Classification Model Based on Deep Residual Shrinkage Network Combined with Dilated Convolution. 12th International Conference on Intelligent Information Processing (IIP), May 2022, Qingdao, China. pp.409-418, 10.1007/978-3-031-03948-5_33 . hal-04178747

HAL Id: hal-04178747

<https://inria.hal.science/hal-04178747v1>

Submitted on 8 Aug 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

A HEP-2 Cell Image Classification Model Based on Deep Residual Shrinkage Network Combined with Dilated Convolution

Wang Chen^[1-2], He Tao², Liu Jiansheng², Li Dapeng², Wen Yingyou^[1-2]

¹ Northeastern University, Shenyang 110004, China

² Neusoft Research, Shenyang, Liaoning 110179, China

wangchen-neu@neusoft.com, het@neusoft.com

Abstract.

Detection of antinuclear antibodies (ANA) in human epithelial cells (HEp-2) is a common method for the diagnosis of autoimmune diseases. The recognition of fluorescence images of human epithelial cells obtained using Indirect Immunofluorescence (IIF) is a key step in the classification of ANA. To address the problems of low efficiency and high labor intensity caused by manual evaluation methods, a HEP-2 cell fluorescent image classification model based on the depth residual shrinkage network combined with dilated convolution is proposed. First, The model accomplish feature extraction by expanding the field of sensation through dilated convolution. Secondly, a 50-layer deep residual network is build, and each residual module embeds a soft threshold learning sub-network to shrink the output data of the original residual module by the automatically learned soft threshold to achieve the purpose of noise removal. Finally, the multi-scale feature fusion module and fusion of features at different scales are used to complete the classification of HEP-2 cell. Experiments show that this model has good performance and is superior to other depth neural network methods.

Keywords: Deep residual shrinkage network, soft threshold, ANA, HEP-2

1 Introduction

Autoimmune disease refers to the immune system mistakenly attacking healthy cells in human tissues. In this case, the body will release various antibodies, one of which is antinuclear antibody (ANA). The specificity of the type of antinuclear antibody varies with different autoimmune diseases. Therefore, the identification of ANA can be used as a means to assist medical treatment and evaluate the degree of damage in patients, which has important medical value. Indirect immunofluorescence (IIF) [1] is the most commonly used method to detect ANAs. IIF imaging has good specificity and sensitivity. Recognition of IIF fluorescence images can classify different antinuclear antibody in patients' serum, it plays an important role in the diagnosis of specific autoimmune diseases.

According to the cell nuclear fluorescence imaging features , ANA are usually divided into 9 categories as shown in Fig.1: (1)cytoplasmic granular type, (2)cytoplasmic fibrous type, (3) nuclear dot type, (4)nuclear granular type, (5)nuclear membrane type, (6) nucleolar type, (7)homogeneous type, (8)centromere type and (9)negative control type. Each type of picture has great similarity in morphology, which makes it difficult to extract picture features.

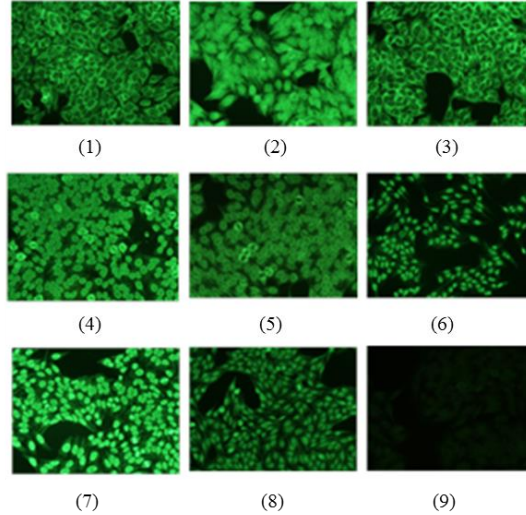


Fig. 1. ANA categories according to the cell nuclear fluorescence imaging features

The traditional HEp-2 cell image recognition[2-3] method not only requires experienced doctors to spend a lot of time on manual discrimination, but also different doctors' judgment methods for cell types are relatively subjective, and different doctors' judgments will be inconsistent[4]. Therefore, building an automatic and reliable HEp-2 image recognition model has become an important research topic.

2 HEp-2 Cell Classification

At present, in-depth research has been carried out in the field of HEp-2 cell classification, especially convolutional neural network is more widely used in this field. Manju C [5] proposed an improved convolutional neural network to effectively classify HEp-2 cells by using the optimization concept. Rahmans [6] reviewed the existing Hep-2 cell image classification methods based on deep learning, and analyzed the core idea, remarkable achievements, key advantages and disadvantages of each method. Foggia [7] comprehensively summarized the development of HEp-2 cell classification methods. Hobson [8] proposed classification methods on large datasets, and analyzed algorithms and their performance at different levels of cells and samples. Manivannan [9] extracted root sieve features and multi-resolution local patterns from HEp-2 cell images and classified them according to the set of support vector machines. Lei [10] proposed a cross modal transfer learning strategy to automatically identify HEp-2

cells. Xie [11] designed a deep monitoring network based on ResNet, which can directly guide the training of the upper and lower layers of the network and offset the impact of unstable gradient changes on the training process. However, most of the existing HEp-2 cell recognition methods are for the processing of single cell samples, and the actual demand is often the whole picture containing many different forms. Therefore, this paper proposes a picture level HEp-2 image recognition model based on deep residual shrinkage network combined with dilated convolution (DRSN-DC). This model uses the ResNet model as the base network, extracts features by expanding the perceptual field through dilated convolution [12], removes noise and redundant data using a deep residual shrinkage network[13], and finally goes through a multi-scale[14] feature fusion module and fuses features from different scales to improve the classification accuracy.

3 HEp-2 Image Classification Model Based on DRSN-DC

3.1 Model Architecture

The HEp-2 classification model proposed in this paper mainly consists of dilated convolution module, depth residual shrinkage module and multi-scale feature fusion module. The dilated convolution module automatically extracts feature information at different scales in the hep-2 cell images by using a dilated convolution network with expansion rates of 1, 2 and 3, respectively, through three convolution operations with different sensory fields. The deep residual shrinkage module removes noise and redundant data from the features extracted by the dilated convolution network through soft thresholding to enhance the weights of effective features and improve the classification accuracy of hep-2 cell images. The multi-scale feature fusion module takes the obtained multi-scale feature information, performs feature fusion, and outputs the classification results. The model architecture is shown in Fig.2.

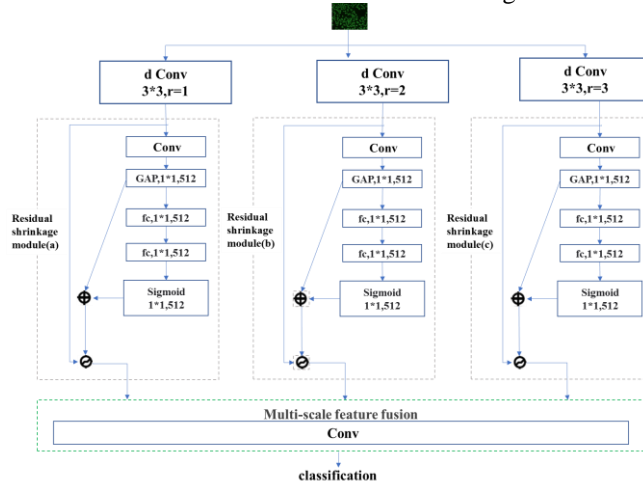


Fig. 2. Model architecture

3.2 Dilated Convolution Network construction

Traditional convolutional neural networks have better performance in image classification, but they also have some drawbacks. For example, although the pooling operation can increase the perceptual field, it can lead to information loss. Dilated convolution defines the spacing of the convolution kernel processing data values during the convolution operation by introducing the expansion rate in the convolution layer, thus increasing the perceptual field, which can improve the segmentation effect of small object recognition in tasks such as target detection and semantic segmentation. HEp-2 cell fluorescence images have the characteristics of small targets, dense distribution and irregular distribution, etc. With the increase of the number of layers in the network, the local information is lost during the down sampling process, resulting in small cells cannot be recognized. The dilated convolution network can increase the perceptual field of feature extraction to capture more detailed features without losing local information. In this paper, three dilated convolutional networks with kernel of 3×3 and dilation rate r of 1, 2, and 3 are used, as shown in Fig.3.

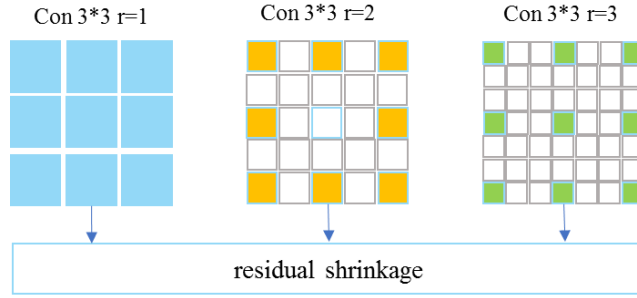


Fig. 3. Three dilated convolutional networks with kernel of 3×3

3.3 Deep Residual Learning Network Construction

Deep residual learning [15] is a very effective method to solve the problem of network degradation. This method does not directly use several layers of networks to fit the desired actual mapping relationship, but to fit a residual mapping. This way makes the output change have a greater effect on the adjustment of weight, so it is easier to train.

This paper uses a 50 layer depth residual network, and its overall network structure is shown in Fig.4. This network includes 3-layer residual learning units, including three residual units that output 256 feature maps, four residual units that output 512 feature maps, six residual units that output 1024 feature maps, and three residual units that output 2048 feature maps. In the middle 3×3 use before and after convolution 1×1 , the purpose is to reduce the dimension and reduce the amount of calculation.

For the layers with the same size of the output feature map, there are the same number of convolution kernels. When the size of the feature map is halved, it is necessary to double the number of filters to maintain the time complexity of each layer. The output

of the network includes a global mean pooling layer and a fully connected layers (FC) containing 1000 neurons, which are classified by softmax.

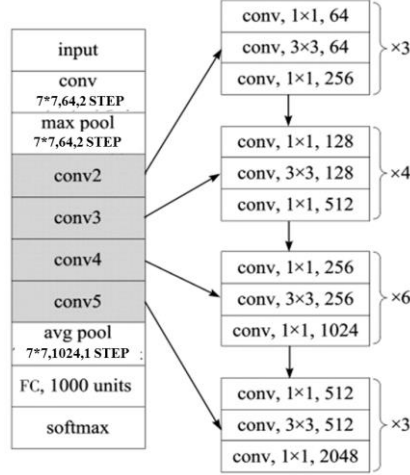


Fig. 4. Depth residual network structure

3.4 Deep Residual Shrinkage Network Construction

In a HEp-2 cell fluorescence image, only part of the cell morphology and connection can reflect the type of the ANA, while other part of the cells cannot be used as the basis for category determination and be seen as noise. How to remove noise and redundant data has become an important problem affecting the performance of the model.

Soft threshold is the most commonly used technology in signal denoising algorithm. It sets the feature of signal smaller than a certain threshold to 0. The traditional method of setting soft threshold needs experts in the business field to select an appropriate soft threshold through continuous debugging according to their professional knowledge. This method has no learning ability and lack of scalability. Therefore, combining soft threshold learning and depth residual network, through automatic learning and adjustment, a way to effectively eliminate noise information and construct high discriminant features is formed, and its transformation formula is as follows:

$$y = \begin{cases} x - \tau & x > \tau \\ 0 & -\tau \leq x \leq \tau \\ x + \tau & x < -\tau \end{cases} \quad (1)$$

Where x is the input characteristic and y is the output characteristic, τ is a positive threshold. After the feature is calculated, the absolute value is less than τ all features will be set to 0, and features with an absolute value greater than τ shrink toward the center, so it is called depth residual shrinkage network. The soft threshold does not set the negative value to 0 like the ReLu activation function, which can retain useful negative value characteristics. A depth residual shrinkage module structure is shown in Fig. 5.

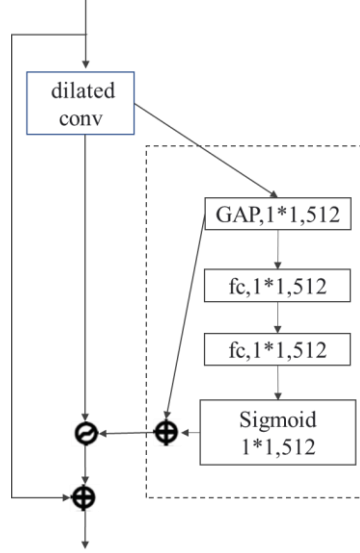


Fig. 5. Depth residual shrinkage module

The depth residual shrinkage network is an improvement of the depth residual network. The soft threshold is inserted into the residual learning unit as a nonlinear transformation layer. The soft threshold learning module is shown in the dotted box in Fig. 5. In this module, firstly, the absolute value of the last layer network output of the residual unit is global average pooling (GAP) to obtain a one-dimensional vector with the same number of convolution kernels as the previous layer. Input the one-dimensional vector into the two-layer fully connected network, apply a sigmoid function at the end of the two-layer FC network, and normalize the scaling parameters.

3.5 Multi-scale feature fusion

The semantic features extracted from small target images in the shallow network are weak. Due to the small and concentrated targets in hEp-2 cell fluorescence images, if the traditional target detection and classification model is used to obtain strong semantic features, the feature maps can only be output at the last layer to complete the classification of the images. However, the last layer is the down sampling process, which will cause the HEP-2 cell fluorescence images with small target to have less effective information on the last feature map, resulting in the loss of local information, thus leading to the degradation of recognition and classification accuracy. Multi-scale feature fusion solves the above problems well. In this paper, the model fuses the features extracted after convolution of three different scales of dilated convolution and features after removing noise and redundant data, and uses Concatenate to stitch the multi-channel extracted feature maps into a complete feature map, and then completes the cell image classification. The multi-scale feature fusion module is shown in Fig.6.

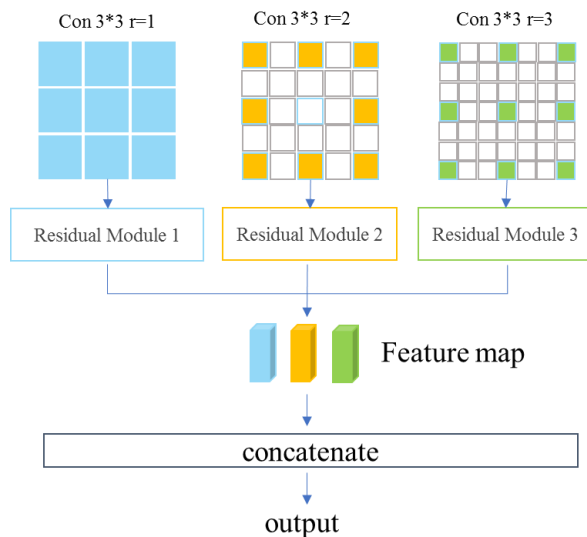


Fig. 6. The multi-scale feature fusion module

4 Test Results and Analysis

The hardware environment for the model experiments in this paper is a Dell R740 server with two NVIDIA Tesla P40 GPU cards mounted; the deep learning framework uses TensorFlow version 1.12.0.

4.1 Experimental Data Set

Due to the cumbersome picture acquisition process, laboratory doctors need to take pictures manually under the microscope, resulting in a scarcity of each type of HEp-2 cell fluorescence image, which is far from meeting the requirements of in-depth learning. It is necessary to enhance the training data set and test data set of HEp-2 cell images through technical means, In order to expand the sample data set and enhance the data characteristics. In order to facilitate training, the original image is cut to 24×224 pixel size.

According to the characteristics of HEp-2 cell fluorescence image imaging and application, rotation and flip are selected for expansion. (1) With the center of the picture as the origin, rotate the picture clockwise for 5 degrees, which can be rotated 72 times, that is, 72 pictures with the same label are derived from one picture; (2) While rotating, flip with the y-axis as the centerline to double the amount of data.

The dataset used in this paper is from the real data of the hospital LIS system, and the annotation and review of the cytofluorescence images are done by professional doctors. Since the number of samples was not sufficient, some methods such as rotation and flip were used to expand the dataset. After removing some unusable data, the

distribution of original image data and expanded image data for each ANA type is shown in Table.1.

Table 1. Distribution of original image data and expanded image data for each ANA type

ANA type	Original image data	Expanded image data
cytoplasmic granular type	208	29952
cytoplasmic fibrous type	183	26352
nuclear dot type	196	28224
nuclear granular type	213	30672
nuclear membrane type	223	32112
nucleolar type	209	30096
homogeneous type	188	27072
centromere type	215	30960
negative control type	220	31680

4.2 Results and Analysis

The error rate is the ratio of the number of misclassified samples to the total number of samples. The curves in Fig.7 show how the overall error rate of the deep learning network constructed in this paper varies with the number of iterations.

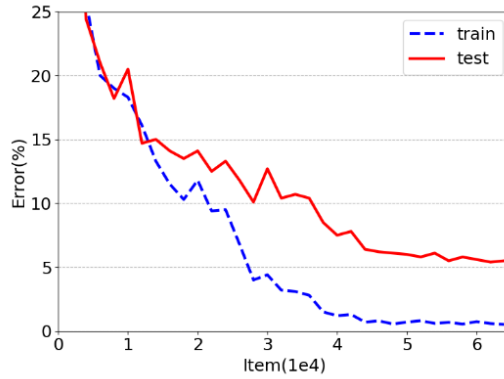


Fig. 7. The error rate of the training set and the test set

From the results, it can be seen that using the deep residual shrinkage network, the error rate of the training set and the test set can decrease quickly with the number of iterations, which better solves the network degradation problem. In this paper, P , R and F_1 score are used to evaluate the performance of the proposed model, and Fig.8 shows the F_1 score for each type of ANA recognition on the test image dataset.

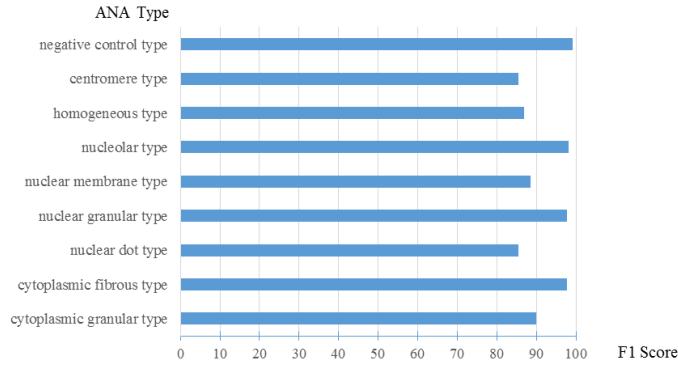


Fig. 8. F_1 score for each type of ANA recognition on the test image dataset

The proposed recognition models in this paper were compared with AlexNet, GoogLeNet, VGGNet and ResNet models on the same dataset for experiments, and the ANA recognition performance of each model is shown in Table.2.

Table 2. ANA recognition performance of each model

Model	P	R	F_1
AlexNet	0.7081	0.7233	0.7156
GoogLeNet	0.8883	0.7759	0.8283
VGG-19	0.7912	0.8331	0.8116
Resnet50	0.8477	0.8699	0.8587
DRSN-DC	0.9310	0.9451	0.9282

By comparing with other models experimentally, the F1 score of the model proposed in this paper reaches 93%, which is about 7% better than the residual network with the same depth, and about 10% better than the F_1 scores of GoogLeNet and VGG-19, showing a significant improvement in performance.

5 Conclusion

In this paper, we propose a HEp-2 cell image classification model based on deep residual shrinkage network combined with dilated convolution. The model embeds soft thresholds as trainable systolic functions into the deep residual network, expands sensory field and accomplish feature extraction by dilated convolution, removes noise and redundant data by using deep residual shrinkage network, and finally fuses features of different scales.. Experiments show that the model in this paper outperforms classical CNN networks such as GoogLeNet and Resnet50 in HEp-2 cell fluorescence image recognition and ANA classification, which greatly improves the speed of clinical fluorescence image analysis and the efficiency of clinical diagnosis for immune diseases.

Acknowledgments

The work is supported by the Shenyang Science and Technology Plan Project (Grant No. RC210469) and in part by Grants from Liaoning Province Key R&D Program Project (Grant Nos. 2019JH2/10100027).

References

1. Creemers C, Guerti K, Spruyt V, et al. HEP-2 cell pattern segmentation for the support of autoimmune disease diagnosis [C]// Proceedings of the 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies. 1001–1015(2011).
2. Jiang X, Percannella G, Vento M. A verification-based multithreshold probing approach to HEP-2 cell segmentation [J]. *Comput Anal Images Patterns*, 5(10):266–276(2015)
3. Percannella G, Soda P, Vento M. A classification-based approach to segment HEP-2 cells[C]// Computer-Based Medical Systems International Symposium on IEEE. 956–955(2012).
4. Cheng C C, Taur J S, Hsieh T Y, et al. Segmentation of anti-nuclear antibody images based on the watershed approach [C] //IEEE Conf Ind Electron. 1695–1700(2010).
5. Manju C C , Jose M V . HEP-2 CELL CLASSIFICATION BY ADAPTIVE CONVOLUTIONAL LAYER BASED CONVOLUTIONAL NEURAL NETWORK[J]. *Biomedical Engineering: Applications, Basis and Communications*, 2019.
6. Rahman S , Wang L , Sun C , et al. Deep Learning Based HEP-2 Image Classification: A Comprehensive Review[J]. *Medical Image Analysis*, 2020, 65:101764.
7. Foggia P, Percannella G, Saggese A, et al. Pattern recognition in stained HEP-2 cells [J]. *Pattern Recognit.* 47(7):2305–2314(2014).
8. Hobson P, Lovell B , Percannella G, et al. Benchmarking human epithelial type 2 interphase cells classification methods on a very large dataset [J]. *Artif Intell Med.* 65(3): 239–250(2015).
9. Manivannan S, Li W, Akbar S, et al. An automated pattern recognition system for classifying indirect immunofluorescence images of HEP-2 cells and specimens [J]. *Pattern Recognit.* 5(51):12–26(2016).
10. Lei Haijun, Han Tao, Huang Weifeng, et al. Cross-Modal Transfer Learning for HEP-2 Cell Classification Based on Deep Residual Network[C]//2017 IEEE International Symposium on Multimedia.465-468(2017).
11. Xie H, He Y, Lei H, et al. Deeply Supervised Residual Network for HEP-2 Cell Classification[C]//2018 24th International Conference on Pattern Recognition (ICPR). Beijing, 699-703(2018).
12. F Yu, Koltun V. Multi-Scale Context Aggregation by Dilated Convolutions [J]. 2016.
13. Zhao M , Zhong S , Fu X , et al. Deep Residual Shrinkage Networks for Fault Diagnosis[J]. *IEEE Transactions on Industrial Informatics*, 2020, PP (99):1-1.
14. P. Wang, L. Song, X. Guo, H. Wang and L. Cui, "A High-Stability Diagnosis Model Based on a Multiscale Feature Fusion Convolutional Neural Network," in *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-9, 2021, Art no. 3522709, doi: 10.1109/TIM.2021.3102745.
15. He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition [J]. *IEEE*, 2016.