A Multi-Feature Fusion Network for Pathological Identification of Tumour Cells

ZHONGDA LU1, JIAMING ZHAO1, YANG SUN1, FENGXIA XU1, and XINGMING MA2

1College of Electrical and Mechanical Engineering, Qiqihar University, Qiqihar, Heilongjiang 161000, China
2State Grid Heilongjiang Provincial Electric Power Co., Ltd. Daqing Power Supply Company, Daqing, Heilongjiang 163458, China

Corresponding author: Fengxia Xu (e-mail: xufengxia_hit@163.com).

This work was supported in part by Science and Technology Project of State Grid Heilongjiang Electric Power Co., Ltd (5224162000JK), in part by Natural Science Foundation of Heilongjiang Province (LH2021F057), and in part by Basic Scientific Research Business Expense Research Project of Heilongjiang Provincial Colleges and Universities (135509309,135409102).

ABSTRACT A novel multi-feature fusion neural network (MFNet) is proposed to address the lack of applicability of existing medical aid diagnostic methods for cross-site and cross-tissue cytopathic lesion screening. MFNet consists of a data-sharing layer, a detailed feature transfer branch, a microscopic identification branch, and an auxiliary loss function. The data-sharing layer converts data images into a feature matrix and extracts detailed elements such as cell morphology, contour, and texture. The microscopic recognition branch obtains multilevel elements by convolving the input elements in stages and fusing them, so that the network can focus on high-level semantic elements such as minor differences of cytopathy. The detail feature transfer branch transfers detail elements across layers and achieves cross-layer fusion with semantic elements. The auxiliary loss function enables the network feature classification capability to be enhanced. MFNet is experimentally compared with AlexNet, VGG-16, ResNet-50, and other models on the tumor cell datasets, and the results show that the proposed method can effectively improve the recognition accuracy.

INDEX TERMS Tumour cells, pathology recognition, convolutional neural network, multi-feature fusion.

I. INTRODUCTION

According to statistics, the prevalence of cancer in China has a trend of increasing year by year, with nearly one million new cases each year [1-2], which seriously threatens people's life and health. In the traditional cancer diagnosis process, the pathological diagnosis is made by doctors with the cooperation of objective factors such as medical images [3] and laboratory results [4-5], which is complex, inefficient, easily fatigued, subjective and uncertain, and to a certain extent subject to misdiagnosis and underdiagnosis. Since the 21st century, machine learning technology, represented by deep learning [6] algorithms, has been rapidly developed, and its use in medically assisted diagnosis, which has a positive effect on the early detection and diagnosis of cancer, and also has excellent potential for application [7-8].

In recent years, many scholars have used the feature of convolutional neural networks [9-12] with solid feature extraction ability to improve the efficiency of cancer diagnosis [13-15] by designing feature representation algorithms to extract elements from pathology images, and then using corresponding classifiers [16-18] to classify and identify the components [19-25], based on which more excellent models have been proposed. Gao et al. [26] extracted three-dimensional texture elements of lung nodules and identified benign and malignant lung nodules using a classifier established by a support vector machine [27] with a model sensitivity of 98%. Rathore S [28] proposed an improved multilevel feature extraction method for the overall elements of pathological images, local elements of image blocks, and epithelial elements, and combined it with a support vector machine classifier to achieve grade classification of colon cancer. Roth et al. [29] used deep learning for medical aid detection to improve the sensitivity of detecting sclerotic metastases, lymph nodes, and colon polyps in CT imaging. Setio [30] used a multi-view convolutional network with multiple 2D CNNs to detect pul-
-monary nodule disease, using a combination of data expansion and Dropout for the study, thus avoiding the phenomenon of overfitting during the model training. Zhang [31] proposed a multi-scale residual convolutional neural network and SVM fusion method for automatic IDC detection, and the model has better accuracy and robustness. The methods as mentioned earlier have achieved good results in medical aid diagnosis. Still, little attention has been paid to generic models for the simultaneous diagnosis of different cancers across sites and tissues.

To address the above problems, this paper analyzes and studies tumor images of lung and colon sites, and proposes a novel tumor cytopathology recognition method that can diagnose tumors of lung and colon sites simultaneously. A multi-feature fusion network is constructed using a deep learning method to extract detailed elements such as morphology, contour, and texture of cells through a data-sharing layer. High-level semantic elements such as cytopathic lesions are obtained by convolutional operations on the input elements in stages. High recognition accuracy was achieved on the cytopathic lesion dataset.

II. APPROACH

The similar color texture of lung adenocarcinoma, lung squamous cell carcinoma, and colon cancer cells lead to the high similarity of inter-class elements, as shown in Figure 1. In contrast, the complexity of cell structure and background can lead to low variability of elements in lung tumors [32-36], so there is a higher probability of misclassification during recognition using convolutional neural networks. Since the same problem of intra- and inter-class variation exists for lesioned cells at different sites, accurate identification and screening of tumor cells require consideration of not only detailed elements (e.g., cell morphology, contour, texture, etc.), but also high-level semantic elements such as slight differences in cells to determine the type of cellular lesion.

In response to the lack of applicability of existing medical aid diagnostic methods for cross-site and cross-tissue cytopathic lesion screening, this paper uses deep learning methods to analyze and study images of lung and colon diseases. It proposes a novel tumor cytopathic recognition method (MFNet) for screening three common and high incidence lung adenocarcinoma, lung squamous cell carcinoma, and colon adenocarcinoma in lung and colon sites. Figure 2 shows the MFNet network architecture. The MFNet structure can be divided into four parts: Data-Sharing Layer (DS), Detailed Feature Transfer Branch (DT), Microscopic Identification Branch (MI), and Auxiliary Loss Function. DS aims to first convert the images into feature matrices in a uniform manner. A constant underlying feature matrix is more conducive to high-level feature information extraction while extracting low-level detail elements such as cell morphology, contour, and texture, which can be ignored for higher-level semantic elements. On the other hand, the shared convolutional layer can reduce the network training parameters exponentially. MI achieves the extraction of elements such as minor differences of cell lesions by the attention module of multi-feature extraction and the residual structure, the network focuses on the high-level semantic elements such as minor differences of cell lesions. DT fuses the detailed components extracted by DS with the semantic elements extracted by MI across layers. Subsequently, after global average pooling, feature fusion is performed by Concat method, feature classification is realized by Softmax function, and tumor cell pathology recognition is achieved by enhancing the network feature classification capability through auxiliary loss function.
FIGURE 2. MFNet network structure. DS represents data-sharing Layer, DT represents Detail Feature Transfer Branch, MI represents Microscopic Identification Branch, Loss represents Loss Function. © represents Concat connection, FC represents fully connected layer. Softmax represents Softmax classifier.

FIGURE 3. DS module structure. Conv represents the convolutional layer, © represents the Concat connection, stride represents step size, ⨁ represents sum operation.

III. NETWORK STRUCTURE

A. DATA-SHARING LAYER (DS)

The Data-Sharing layer (DS) has two main roles in the MFNet network. First: Since lung adenocarcinoma, lung squamous cell carcinoma, and colon cancer cell images are characterized by high inter-class feature similarity and low intra-class feature variability, which can cause overfitting or bias in the model, the convolutional layers close to the input is used to extract the detailed elements of the cells using a DS module with more vital characterization ability, but the higher-level semantic features in the images can be ignored. Secondly, the data processed by DS can be shared by the next network structure in MFNet network, such as Microscopic Identification Branch (MI) and Detailed Feature Transfer Branch (DT), and the two branches share the same weights, which can effectively save the computational cost of the model. The structure of the DS network is shown in Figure 3. When the cell images are input to the system, the detailed features of the cell images are extracted through the DS. The DS module contains 2 convolutional layers and 3 convolutional units. Both convolution layer 1 and convolution layer 2 are designed with a 3*3 convolution, which can divide the input feature map into 2 feature submaps. Convolution unit 1 extracts the features by designing 2 branches with different convolutional structures: one branch consisting of 1*1 convolution and 3*3 convolution, and the other branch consisting of 3*3 maximum pooling. 2 different downsampling methods are used to narrow down the feature representation, and the texture features of the cell images can be preserved from being destroyed while the features are effectively extracted for subsequent feature extraction. The output elements of the dual channels are then fused, making the structure computationally efficient and effective in feature representation. Since the extraction of cancer cells requires
different sensory fields to capture detailed elements, the convolution unit 2 adopts a similar two-branch scheme to extract detailed elements by a different number of convolutions on each branch to obtain various sensory fields, while selecting a convolution with a step size of 2 to achieve the reduction of the data size. Convolution unit 3 adopts residual structure to solve the problem of network degradation. The residual design directly merges the input and output of the module to achieve cross-layer connection, which enables the deep network to learn effectively while effectively preventing gradient explosion. After the DS, the data images are converted into a feature matrix. Meanwhile, the output data features go directly to the MI and the DT, which facilitate the subsequent extraction of high-level feature information by MI and DT. Such a design can effectively extract feature information such as texture and color, and further enhance the model feature expression capability.

B. MICROSCOPIC IDENTIFICATION BRANCH (MI)

1) ATTENTION MODULE FOR MULTI-FEATURE EXTRACTION (MF MODULE)

For feature extraction of lesioned cells, it is necessary to pay attention to not only detailed elements but also advanced semantic elements such as minor differences in cell lesions, so the attention module (MF module) for multi-feature extraction is designed.

The MF module consists of a multi-feature extraction module and an attention module [37], and the structure is shown in Figure 4. When the feature information is input, Step 1, the input feature information is divided into two equal parts equally, half of the feature information is subjected to 1*1 convolution operation and sent to the end for merging, while the other half of the feature information is subjected to convolution operation; Step 2, the feature information after the convolution operation in Step 1 is taken as the input and subjected to average partitioning, half of the feature information is subjected to convolution operation, and the other half of the feature information is directly Step 3, repeat the above steps until the last feature information; Step 4, merge the previous feature information with the previously obtained feature information, so that different groups of feature information can enjoy various perceptual fields, while locating the cytopathic region in the image to get more effective cytopathic elements, and also to alleviate the adverse effects caused by the image size change and increase the richness of feature information. However, the elements extracted from different receptive fields are too different, the attention module is used to fuse the depth and spatial information through local receptive fields to make the overall...
feature expression of the network enhanced.

2) MICROSCOPIC RECOGNITION BRANCH (MI) STRUCTURE

The role of the Microscopic Identification Branch (MI) is to extract deep, high-level features of the image. Since the disparity between pathologies in different parts of human cells is contained in only minute local details, while the feature maps near the output are rich in semantic information, the MI module is used in the deeper layers of the network to extract high-level semantic features that are not extracted by DS. The MI network structure is shown in Figure 5. The MI structure contains two double-branch convolutional units, each containing an attention module "MF module" for multi-feature extraction. When the distribution of lesion information in the image is relatively scattered, MI can obtain more local elements through the attention module of multi-feature extraction. When the distribution of lesion information is relatively concentrated, more global elements can be obtained through 3*3 and 1*1 convolution. Meanwhile, the residual structure is designed in the convolution unit 2, which can effectively mitigate the gradient disappearance, so that both the low-grade texture information is retained and the high-grade semantic information of the convolutional neural network is extracted. Finally, the fusion operation makes the extracted cell image features richer and improves the feature extraction capability of the overall network.

C. DETAILED FEATURE TRANSFER BRANCHING (DT)

In convolutional network design, often more profound and more complex networks can achieve better results and accuracy, but the weak performance improvement brought by blindly stacking convolutional layers and increasing complexity is not suitable for application in practical tasks. Therefore, DT directly fuses the DS extracted data elements with the MI output data elements across layers, which can effectively prevent the problem that the extracted detailed elements are reduced in size and essential local details is lost due to too many convolution operations.

D. AUXILIARY LOSS FUNCTION

To avoid the problem of gradient disappearance when the network is deeper, the network is designed with an auxiliary loss function in DS that conducts the gradient forward and enhances the model feature classification capability. The output of DS is fully connected and classified, and the activation function of the fully connected layer is chosen as RELU, which can make the network training faster and increase the nonlinearity of the network. The classification layer uses Softmax, a function that sums the neuron outputs to 1 and effectively compresses the classification results to 0~1, referring to (1).

\[
\sigma(z) = \frac{e^{z_j}}{\sum_{k=1}^{n} e^{z_k}} \quad (1)
\]

In Eq. (1) \(z_j^l\) is the weighted input of the \(j^{th}\) neuron in the \(l^{th}\) layer. The model optimization loss function is chosen to be the cross-entropy loss function (CELF), referring to (2).

\[
C(\omega, b) = -\frac{1}{n} \sum_{x} [y \ln a + (1 - y) \ln(1 - a)] \quad (2)
\]

In Eq. (2) \(\omega\) represents the set of weights in the model; \(b\) represents the bias; \(n\) represents the number of input training datasets; \(a\) represents the model prediction results; summation is performed over all training inputs \(x\); \(y\) represents the labels with training datasets; and \(C\) represents the loss function output.

Subsequently, the auxiliary loss function output from DS and the loss function output from MI are subjected to weight matching operation to obtain the total network loss, referring to (3). The use of the auxiliary loss not only increases the back propagation of the loss function, but also acts as a regular constraint term, which enhances the feature classification ability of the network and thus improves the model accuracy. During the training process, the auxiliary loss is added to the total loss of the network in the form of weight shares.

\[
Q = \alpha C(\omega, b) + \beta C(\omega, b) \quad (3)
\]

In equation (3), \(Q\) is the total loss function output, \(C\) represents the loss function, \(w\) represents the set of weights in the model; \(b\) represents the bias; \(\alpha\) and \(\beta\) are the auxiliary loss function for the DS output and the loss function matching weights for the MI output, respectively.

IV. EXPERIMENTS

A. EXPERIMENTAL PREPARATION

The experiment was conducted using the open source cytopathic dataset of lung and colon sites shown in Figure 1. The dataset contains an image dataset of 25,000 color images from 5 categories (LC25000). Each category contains 5000 images of the following histological entities: colon adenocarcinoma, benign colon tissue, lung adenocarcinoma, lung squamous cell carcinoma and benign lung tissue. The ratio of the training set to the test set was 7:3, and the same data distribution was maintained between the training and test sets, as shown in Table 1, with 17,500 images in the training set and 7,500 images in the validation set. The specific configuration of the experimental environment is shown in Table 2, with a 32G graphics memory NVIDIA Tesla V100 graphics card and a deep learning framework using PaddlePaddle version 1.8.4.

| Lung, colon area | Data sets | Training set | Validation set |
|------------------|-----------|-------------|----------------|
| Cytopathic conditions |          |             |                |
| Adenocarcinoma of the colon | 5000 | 3500 | 1500 |
| Benign colonic tissue | 5000 | 3500 | 1500 |
| Adenocarcinoma of the lung | 5000 | 3500 | 1500 |
| Squamous cell carcinoma of the lung | 5000 | 3500 | 1500 |
| Benign lung tissue | 5000 | 3500 | 1500 |
TABLE 2. Experimental environment configuration table

| Experimental environment configuration items | Parameter values |
|---------------------------------------------|-------------------|
| CPU                                         | Inter(R) Xeon(R) Gold 6271C @2.60GHz |
| Video Cards                                 | NVIDIA Tesla V100 GPU 32G |
| Memory                                      | 32G |
| Disks                                       | 100G |
| Deep learning frameworks                    | PaddlePaddle 1.8.4 |
| Operating systems                           | Ubuntu 16.04.1 LTS (64-bits) |
| Other tools                                 | Python 3.7.1 CUDA10.1 |

TABLE 3. Table of experimental parameters

| Parameter name     | Parameter values |
|--------------------|-------------------|
| Image size         | 128*128           |
| Batch size         | 512               |
| Learning rate      | 0.0001            |
| Training period    | 60                |

The experimental parameters were set as shown in Table 3. The input image size was 128*128, the GPU was used to speed up the neural network training, the adaptive moment estimation (Adam) was chosen as the optimizer of the model parameters, the learning rate was 0.0001, and the loss function Cross entropy loss was used, the neural network was trained with 512 images per batchsize, and the training period was 60 epochs.

B. COMPARATIVE EXPERIMENTS

For the effectiveness of the MFNet method, comparative analysis experiments were conducted between MFNet and traditional CNN models: AlexNet [38], VGG-16 [39], and ResNet-50 [40]. The models mentioned above were trained separately from initial to convergence on the cytopathic datasets of lung and colon sites, ensuring training from scratch under the same conditions. In order to be able to show more intuitively the performance of the models during training, the strategy of testing the validation set immediately after each training iteration of the training set was used in the model training process. Figure 6 shows the variation of accuracy and loss function for different CNN models on the cytopathic dataset of lung and colon sites.

Figure 6 shows the variation of accuracy and loss function of different CNN models on the cytopathic datasets of lung and colon sites. In the Figure, MFNet has a higher accuracy.

The input image size was 128*128, the GPU was used to speed up the neural network training, the adaptive moment estimation (Adam) was chosen as the optimizer of the model parameters, the learning rate was 0.0001, and the loss function Cross entropy loss was used, the neural network was trained with 512 images per batchsize, and the training period was 60 epochs.

TABLE 4. Accuracy of each model

| Network Model | Recognition accuracy (Accuracy) |
|---------------|---------------------------------|
| AlexNet       | 89.013%                         |
| VGG-16        | 96.046%                         |
| ResNet-50     | 97.389%                         |
| MFNet         | 98.707%                         |

![Figure 6. Training results of different network models.](image)
rate than the above models. The training results of the lung and colon cytoplasm dataset showed that MFNet achieved higher accuracy without affecting the drop rate and convergence rate of the loss function of the model, while not causing serious overfitting due to the model own structure. The model training results for MFNet, AlexNet, VGG-16, and ResNet-50 are shown in Table 4. Figure 7 shows the variation of accuracy rates of different models on the cytopathic dataset of lung and colon sites, with an overall increasing and gradual stabilization of each model in the Figure.

As shown in Figure 6, the AlexNet, VGG-16, and ResNet-50 models all reached convergence within 60 iteration cycles, and MFNet has higher accuracy than the above models. MFNet have a multi-feature extraction attention module that can extract data elements more effectively. The training results from the cytopathic dataset of lung and colon sites show that MFNet achieves higher accuracy without affecting the drop rate and convergence rate of the loss function of the model, while no serious overfitting phenomenon occurs due to the model's own structure. The model training results of MFNet, AlexNet, VGG-16, and ResNet-50 are shown in Table 4. Figure 7 shows the variation of accuracy of different models on the cytopathic dataset of lung and colon sites, the overall trend of each model in the figure is increasing and gradually stabilizing.

C. ABLATION EXPERIMENT

To verify the effectiveness of MFNet detail feature transmission branch (DT), the DT of MFNet was removed to obtain ND-MFNet. Figure 8 (a) shows the change curves of MFNet and ND-MFNet accuracy. The accuracy of ND-MFNet finally converges to 97.095%, and the accuracy of MFNet finally converges to 98.707%, which is 1.612% higher than that of ND-MFNet. The experimental results show that the detail feature transmission branch (DT) of MFNet can effectively extract the detail elements such as morphology, contour and texture of cells, which has a certain contribution to the accuracy improvement of the model.

To verify the effectiveness of the attention module (MF module) of MFNet multi-feature extraction, the MF module in the microscopic discrimination branch (MI) of MFNet was
removed to obtain NA-MFNet. Figure 8 (b) shows the accuracy change curves of MFNet and NA-MFNet. As can be seen from Figure 8 (b), both NA-MFNet and MFNet accuracies showed an increasing trend, and the accuracy of NA-MFNet increased significantly after 35 rounds. The accuracy of NA-MFNet finally converged at 96.153%, and the accuracy of MFNet was finally 2.554% higher than that of NA-MFNet. As the MF module enables different groups of feature information to enjoy different perceptual fields, thus increasing the richness of feature information, and the fusion of elements through the channel attention module, it enables the extraction of advanced semantic elements such as minor differences in cell lesions by MI, increasing the diversity of captured elements and effectively improving the accuracy of the model.

To verify the effectiveness of the MFNet auxiliary loss function, the NW-MFNet was obtained by removing the MFNet auxiliary loss function. Figure 8 (c) shows the change curves of MFNet and NW-MFNet accuracy. As can be seen from Figure 8 (c), the accuracy of ND-MFNet and MFNet increased alternately in the first 35 rounds, and the accuracy of MFNet increased significantly compared with that of ND-MFNet in the last 25 rounds and gradually stabilized. The accuracy of NW-MFNet finally converged to 97.021%, the accuracy of MFNet was finally 1.686% higher than that of NW-MFNet, indicating that the auxiliary loss function can not only increase the back propagation of the loss function, but also act as a regular constraint term, which can effectively enhance the feature classification ability of the network and make the model accuracy significantly better.

The recognition of different cell lesions by the four model structures of MFNet, ND-MFNet, NA-MFNet, and NW-MFNet is shown in Table 5. The recognition accuracy of MFNet is higher than the other four models for three types of lesion cells and two types of benign cell tissues, indicating that the DT of MFNet can effectively fuse the morphology, contour, texture, and other detailed elements of the DS-extracted lesion cells with high-level semantic elements. The MF module can effectively extract high-level semantic elements such as minor differences in cell lesions, and the auxiliary loss function can effectively enhance the feature classification ability of the network. Impurities in the feature information can be effectively removed. Therefore, this paper finally combines DT, MF module and auxiliary loss function to form MFNet, which can extract multi-feature data elements while alleviating the adverse effects caused by the similarity of cellular lesions, thus enhancing the expression ability of the model and improving the model accuracy.

To validate the recognition performance of MFNet, it was compared with other medical aid detection models, as shown in Table 6, and its accuracy was improved by 10.447%, 3.707%, 8.857%, 3.707% and 2.047% respectively compared with the other five models. The results show that the MFNet model proposed in this paper is advanced in the identification of cytopathic lesions in the lung and colon sites.

To verify the influence of different weight ratios of the auxiliary loss function on the network accuracy, the experiment took nine groups of weight ratios than batch experiments, and the weight ratios experiment is shown in Figure 9, the experimental results are shown in Table 7, cost1 represents the MI output loss function, cost2 represents the DS output auxiliary loss function. As shown in Figure 9, different weight ratios lead to varying degrees of feature extraction and thus have a certain impact on the network accuracy. By adjusting the weight ratios, the feature classification capability of the network is enhanced, and the model recognition accuracy is further improved. From Table

### TABLE 5. Ablation experiment

| Structure                  | MFNet | ND-MFNet | NA-MFNet | NW-MFNet |
|---------------------------|-------|----------|----------|----------|
| Adenocarcinoma of the lung| 99.39%| 93.913%  | 93.09%   | 92.972%  |
| Benign lung tissue        | 100%  | 99.931%  | 100%     | 100%     |
| Squamous cell carcinoma of the lung | 99.14%| 96.104%  | 93.582%  | 94.754%  |
| Adenocarcinoma of the colon | 99.672%| 99.672%  | 99.476%  | 99.608%  |
| Benign colonic tissue     | 100%  | 99.541%  | 99.672%  | 99.74%   |
| Cytopathic data on lung and colon sites | 98.707%| 97.095%  | 96.153%  | 97.021%  |

### TABLE 6. Model comparison experiments

| Network Model | Accuracy (Accuracy) |
|---------------|---------------------|
| Literature [31] | 88.26%              |
| Literature [41] | 95%                 |
| Literature [42] | 89.85%              |
| Literature [43] | 95%                 |
| Literature [44] | 96.66%              |
| MFNet         | 98.707%             |

### TABLE 7. Experimental results for weighting ratio groups

| Cost | Acc (%) |
|------|---------|
| 0.1  | 98.252  |
| 0.2  | 98.350  |
| 0.3  | 98.182  |
| 0.4  | 97.868  |
| 0.5  | 97.429  |
| 0.6  | 97.333  |
| 0.7  | 97.077  |
| 0.8  | 97.964  |
| 0.9  | 96.833  |

![FIGURE 9. Weighting ratio experiment](image-url)

The accuracy change curves of MFNet and NW-MFNet. As can be seen from Figure 8 (b), both NA-MFNet and MFNet accuracies showed an increasing trend, and the accuracy of NA-MFNet increased significantly after 35 rounds. The accuracy of NA-MFNet finally converged at 96.153%, and the accuracy of MFNet was finally 2.554% higher than that of NA-MFNet. As the MF module enables different groups of feature information to enjoy different perceptual fields, thus increasing the richness of feature information, and the fusion of elements through the channel attention module, it enables the extraction of advanced semantic elements such as minor differences in cell lesions by MI, increasing the diversity of captured elements and effectively improving the accuracy of the model.

To verify the effectiveness of the MFNet auxiliary loss function, the NW-MFNet was obtained by removing the MFNet auxiliary loss function. Figure 8 (c) shows the change curves of MFNet and NW-MFNet accuracy. As can be seen from Figure 8 (c), the accuracy of ND-MFNet and MFNet increased alternately in the first 35 rounds, and the accuracy of MFNet increased significantly compared with that of ND-MFNet in the last 25 rounds and gradually stabilized. The accuracy of NW-MFNet finally converged to 97.021%, the accuracy of MFNet was finally 1.686% higher than that of NW-MFNet, indicating that the auxiliary loss function can not only increase the back propagation of the loss function, but also act as a regular constraint term, which can effectively enhance the feature classification ability of the network and make the model accuracy significantly better.

The recognition of different cell lesions by the four model structures of MFNet, ND-MFNet, NA-MFNet, and NW-MFNet is shown in Table 5. The recognition accuracy of MFNet is higher than the other four models for three types of lesion cells and two types of benign cell tissues, indicating that the DT of MFNet can effectively fuse the morphology, contour, texture, and other detailed elements of the DS-extracted lesion cells with high-level semantic elements. The MF module can effectively extract high-level semantic elements such as minor differences in cell lesions, and the auxiliary loss function can effectively enhance the feature classification ability of the network. Impurities in the feature information can be effectively removed. Therefore, this paper finally combines DT, MF module and auxiliary loss function to form MFNet, which can extract multi-feature data elements while alleviating the adverse effects caused by the similarity of cellular lesions, thus enhancing the expression ability of the model and improving the model accuracy.

To validate the recognition performance of MFNet, it was compared with other medical aid detection models, as shown in Table 6, and its accuracy was improved by 10.447%, 3.707%, 8.857%, 3.707% and 2.047% respectively compared with the other five models. The results show that the MFNet model proposed in this paper is advanced in the identification of cytopathic lesions in the lung and colon sites.

To verify the influence of different weight ratios of the auxiliary loss function on the network accuracy, the experiment took nine groups of weight ratios than batch experiments, and the weight ratios experiment is shown in Figure 9, the experimental results are shown in Table 7, cost1 represents the MI output loss function, cost2 represents the DS output auxiliary loss function. As shown in Figure 9, different weight ratios lead to varying degrees of feature extraction and thus have a certain impact on the network accuracy. By adjusting the weight ratios, the feature classification capability of the network is enhanced, and the model recognition accuracy is further improved. From Table
9, it can be seen that the model has the highest accuracy when the weighting ratio of cost1 and cost2 is 4:6, so cost1:cost2=4:6 is selected as the weighting ratio of MFNet.

V. CONCLUSION

In this paper, we propose a novel multi-feature fusion method for tumor cytopathology recognition, which can effectively extract detailed elements and semantic elements of cytopathic images, alleviate the adverse effects caused by cytopathic similarity, and enhance the expression capability of cytopathic images, alleviate the adverse effects caused by cytopathic similarity, and enhance the expression capability of cytopathic images, alleviate the adverse effects caused by cytopathic similarity, and enhance the expression capability of cytopathic images, alleviate the adverse effects caused by cytopathic similarity, and enhance the expression capability of cytopathic images.

REFERENCES

[1] R. L. Siegel, K. D. Miller, and A. Jemal, “Cancer statistics, 2020,” CA A Cancer J Clin, vol. 70, no. 1, pp. 7 – 30, Jan. 2020.

[2] K. SUN, R. ZHENH, S. ZHANG et al., “Progress of Cancer Screening Program in Urban China,” China Cancer, Vol. 28, no.1, pp.1-11, 2019.

[3] G. Litjens et al., “A survey on deep learning in medical image analysis,” Medical Image Analysis, vol. 42, pp. 60 – 88, Dec. 2017.

[4] Y. Shin, J. K. Yoon, and G. Marwaha, “Progress in the Treatment and Outcomes for Early-Stage Non-minor Cell Lung Cancer,” Lung, vol. 196, no. 3, pp. 351 – 358, Jun. 2018.

[5] J. W. Clark and D. L. Longo, “Recent progress in systemic treatment for lung cancer,” Current Opinion in Pulmonary Medicine, vol. 24, no. 4, pp. 355 – 366, Jul. 2018.

[6] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436-436, May. 2015.

[7] A. Srivastava, C. Kulkarni, K. Huang, A. Parwani, P. Malikic, R. Machiraju, “Imitating Pathologist Based Assessment With Interpretable and Context Based Neural Network Modeling of Histology Images,” Biomed Inform Insights, vol. 10, p. 117822261880748, Jan. 2018.

[8] B. Ehteshami Bejnordi et al., “Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer,” JAMA, vol. 318, no. 22, p. 2199, Dec. 2017.

[9] F. A. Spanhol, L. S. Oliveira, C. Petjeitain, and L. Heute, “Breast cancer histopathological image classification using convolutional neural networks,” in Proc. Int. Joint Conf. Neural Netw. (IJCNN), Jul. 2016, pp. 2560 – 2567.

[10] A. Bayramoglu, J. Kannala, and J. Heikila, “Deep learning for magnification-independent breast cancer histopathology image classification,” in Proc. 23rd Int. Conf. Pattern Recognit. (ICPR), Dec. 2016, pp. 2440 – 2445.

[11] Q. Li and W. Li, “Using deep learning for breast cancer diagnosis,” Chinese Univ., Hong Kong, Tech. Rep., 2017.

[12] Z. Han, B. Wei, Y. Zheng, Y. Yin, K. Li, and S. Li, “Breast cancer multiclassification from histopathological images with structured deep learning model,” Sci. Rep., vol. 7, no. 1, p. 4172, Dec. 2017.

[13] A. Madabhushi and G. Lee, “Image analysis and machine learning in digital pathology: Challenges and opportunities,” Medical Image Analysis, vol. 33, pp. 170 – 175, Oct. 2016.

[14] D. Komura and S. Ishikawa, “Machine Learning Methods for Histopathological Image Analysis,” Computational and Structural Biotechnology Journal, vol. 16, pp. 34 – 42, 2018.

[15] E. Ozdemir and C. Gunduz-Demir, “A Hybrid Classification Model for Digital Pathology Using Structural and Statistical Pattern Recognition,” IEEE Trans. Med. Imaging, vol. 32, no. 2, pp. 474 – 483, Feb. 2013.

[16] B. E. Bejnordi, G. Litjens, N. Timofeeva, I. Otte-Holler, A. Homeyer, N. Karssenmeijer, and J. A. van der Laak, “Stain specific standardization of whole-slide histopathological images,” IEEE Trans. Med. Imag., vol. 35, no. 2, pp. 404 – 415, Feb. 2016.

[17] R. Nateghi, H. Danyali, and M. S. Helfroush, “Maximized inter-class weighted mean for fast and accurate mitosis cells detection in breast cancer histopathological images,” J. Med. Syst., vol. 41, no. 9, p. 146, Sep. 2017.

[18] T. H. Vu, H. S. Mousavi, V. Monga, G. Rao, and U. K. A. Rao, “Histopathological image logical classification using discriminative feature-oriented dictionary learning,” IEEE Trans. Med. Imag., vol. 35, no. 3, pp. 738 – 751, Mar. 2016.

[19] A. F. M. Agarap, “A neural network architecture combining gated recurrent unit (GRU) and support vector machine (SVM) for intrusion detection in network traffic data,” in Proc. 10th Int. Conf. Mach. Learn. Comput., Feb. 2018, pp. 26 – 30.

[20] D. Xue, R. Zhang, H. Feng, and Y.-L. Wang, “CNN-SVM for microvascular morphological type recognition with data augmentation,” J. Med. Biol. Eng., vol. 36, no. 6, pp. 755 – 764, Dec. 2016.

[21] A. S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, “CNN elements off-the-shelf: An astounding baseline for recognition,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops, Jun. 2014, pp. 512 – 519.

[22] G. Olgun, C. Sokmensuer, and C. Gunduz-Demir, “Local object patterns for the representation and classification of colon tissue images,” IEEE J. Biomed. Health Inform., vol. 18, no. 4, pp. 1390 – 1396, Jul. 2014.

[23] Y. S. Vang, Z. Chen, and X. Xie, “Deep learning framework for multiclass breast cancer histology image classification,” in Image Analysis and Recognition, Lecture Notes in Computer Science, vol. 10882, New York, NY, USA: Springer, 2018, pp. 914 – 922.

[24] S. Vesal, N. Raviikumar, A. Davari, S. Ellmann, and A. Maier, “Classification of breast cancer histology images using transfer learning,” in Image Analysis and Recognition (Lecture Notes in Computer Science), vol. 10882, New York, NY, USA: Springer, 2018, pp. 763 – 770.

[25] Y. Liu et al., “Detecting Cancer Metastases on Gigapixel Pathology Images,” arXiv:1703.02442 [cs], Mar. 2017, Accessed: Nov. 19, 2021. [Online].

[26] N. Gao et al., “Three-Dimensional Texture Feature Analysis of Pulmonary Nodules in CT Images: Lung Cancer Predictive Models Based on Support Vector Machine Classifier,” J Digit Imaging, vol. 33, no. 2, pp. 414 – 422, Apr. 2020.

[27] C. Cortes and V. Vapnik, “Support-vector networks,” Mach. Learn., vol. 20, no. 3, pp. 273 – 297, 1995.

[28] S. Rathore, M. A. Iftikhar, A. Chaddad, T. Niazi, T. Karasic, and M. Bilello, “Segmentation and Grade Prediction of Colon Cancer Digital Pathology Images Across Multiple Institutions,” Cancers, vol. 11, no. 11, p. 1700, Nov. 2019.

[29] H. R. Roth et al., “Improving Computer-Aided Detection Using Convolutional Neural Networks and Random View Aggregation,” IEEE Trans. Med. Imaging, vol. 35, no. 5, pp. 1170 – 1181, May 2016.

[30] A. A. A. Setio et al., “Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Networks,” IEEE Trans. Med. Imaging, vol. 35, no. 5, pp. 1160 – 1169, May 2016.
Z.Lu et al.: A Multi-Feature Fusion Network for Pathological Identification of Tumour Cells

[31] J. Zhang, X. Guo, B. Wang, and W. Cui, “Automatic Detection of Invasive Ductal Carcinoma Based on the Fusion of Multi-Scale Residual Convolutional Neural Network and SVM,” IEEE Access, vol. 9, pp. 40308 - 40317, 2021.

[32] X. D. Teng, “World Health Organization classification of tumours, pathology and genetics of tumours of the lung,” Zhonghua Bing Li Xue Za Zhi/Chin. J. Pathol., vol. 34, no. 8, p. 544, 2005.

[33] T. Yoshiya, T. Mimae, and Y. Tsutani, “Prognostic role of subtype classification in minor-sized pathologic N0 invasive lung adenocarcinoma,” Ann. Thoracic Surg., vol. 102, no. 5, pp. 1668 – 1673, 2016.

[34] A. Warth, T. Muley, M. Meister, A. Stenzinger, M. Thomas, P. Schirrmacher, P. A. Schnabel, J. Budczies, H. Hoffmann, and W. Weichert, “The novel histologic international association for the study of lung cancer/American thoracic society/European respiratory society classification system of lung adenocarcinoma is a stage independent predictor of survival,” J. Clin. Oncol., vol. 30, no. 13, pp. 1438 – 1446, May 2012.

[35] S. Kundu, R. Mitra, and S. Misra, “Squamous cell carcinoma lung with progressive systemic sclerosis,” J. Assoc. Phys. India, vol. 60, no. 12, pp. 52 – 54, 2012.

[36] S. Davis, K. E. Stanley, R. Yesner, D. T. Kuang, and J. F. Morris, “minor-cell carcinoma of the lung – survival according to histologic subtype: A veterans administration lung group study,” Cancer, vol. 47, no. 7, pp. 1863 – 1866, Apr. 1981.

[37] J. Hu, L. Shen, and G. Sun, “Squeeze-and-excitation networks,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Jun. 2018, pp. 7132 – 7141.

[38] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” Commun. ACM, vol. 60, no. 6, pp. 84 – 90, May 2017.

[39] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” Comput. Sci., 2014.

[40] K. He, X. Zhang, and S. Ren, “Deep residual learning for image recognition,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2016, pp. 770 – 778.

[41] W. Wang, Y. Hu, T. Zou, H. Liu, J. Wang, and X. Wang, “A New Image Classification Approach via Improved MobileNet Models with Local Receptive Field Expansion in Shallow Layers,” Computational Intelligence and Neuroscience, vol. 2020, pp. 1 – 10, Aug. 2020.

[42] S. Pang, Y. Zhang, M. Ding, X. Wang, and X. Xie, “A Deep Model for Lung Cancer Type Identification by Densely Connected Convolutional Networks and Adaptive Boosting,” IEEE Access, vol. 8, pp. 4799 – 4805, 2020.

[43] P. Kainz, M. Pfeiffer, and M. Urschler, “Segmentation and classification of colon glands with deep convolutional neural networks and total variation regularization,” PeerJ, vol. 5, p. e3874, Oct. 2017.

[44] W. J. Sori, J. Feng, A. W. Godana, S. Liu, and D. J. Gelmecha, “DFD-Net: lung cancer detection from denoised CT scan image using deep learning,” Front. Comput. Sci., vol. 15, no. 2, p. 152701, Apr. 2021, doi: 10.1007/s11704-020-9050-z.