A computer-aid multi-task light-weight network for macroscopic feces diagnosis

Ziyuan Yang1,2 · Lu Leng1,3 · Ming Li4 · Jun Chu1

Received: 18 January 2021 / Revised: 15 June 2021 / Accepted: 31 January 2022
Published online: 28 February 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
The abnormal traits and colors of feces typically indicate that the patients are probably suffering from tumor or digestive-system diseases. Thus a fast, accurate and automatic health diagnosis system based on feces is urgently necessary for improving the examination speed and reducing the infection risk. The rarity of the pathological images would deteriorate the accuracy performance of the trained models. In order to alleviate this problem, we employ augmentation and over-sampling to expand the samples of the classes that have few samples in the training batch. In order to achieve an impressive recognition performance and leverage the latent correlation between the traits and colors of feces pathological samples, a multi-task network is developed to recognize colors and traits of the macroscopic feces images. The parameter number of a single multi-task network is generally much smaller than the total parameter number of multiple single-task networks, so the storage cost is reduced. The loss function of the multi-task network is the weighted sum of the losses of the two tasks. In this paper, the weights of the tasks are determined according to their difficulty levels that are measured by the fitted linear functions. The sufficient experiments confirm that the proposed method can yield higher accuracies, and the efficiency is also improved.

Keywords Multi-task diagnosing · Data augmentation, over-sampling · Weighting selection · Macroscopic feces image · Color recognition · Trait recognition
1 Introduction

The health condition of digestive system could be reflected from the macroscopic feces images, which could help preliminarily diagnose the diseases early. A large number of medical staffs are required to do the feces examination, which leads to a high infection risk, and the terrible smell would make medical staffs unable to work efficiently for a long time. The abnormal traits and colors of feces typically indicate that the patients are probably suffering from tumor or digestive-system diseases [8]. In addition, feces can also be applied to detect other diseases, such as COVID-19, and assess prognosis [36]. There are three main advantages for a computer-aid diagnosing approach. 1. There is no need to ask medical staffs to examine the feces in a manual way, which greatly reduces the infection risk and workload; 2. The automatic recognition is fast and the results are reliable; 3 The machine can stably work for a long time even with high-intensity workload.

Feces examinations include macroscopic examination and microscopic examination. In recent years, researchers proposed different methods for microscopic image-based diagnosing. Du et al. [16] proposed a novel framework, which is composed of convolutional neural network (CNN) and principal component analysis, to detect cells in the microscopic fecal images. Li et al. [32] proposed a deep learning-based visible component detection approach for microscopic fecal images. Microscopic examination with microscopes is accurate but time-consuming and expensive, and the laboratorians must have proficient professional skills.

A reliable automatic macroscopic examination is necessary for assessing the health condition of digestive system. Computer technologies have been used to produce gastrointestinal surgery robotics that are deemed as an important next-generation operation mode [13, 27]. Unfortunately, automatic machine is still absent in macroscopic feces examination, so we develop an automatic macroscopic feces examination machine. In the customized medical equipment, the feces are required to put into the tube, and the trait and color of the feces are automatically recognized by computer vision technology. Some feces images are shown in Fig. 1.

To develop a computer-aid multi-task light-weight network for macroscopic feces diagnosis, the contributions of this paper are summarized as follows.

(1) A database labeled in color and trait

![Feces samples in the tubes](image_url)
To our best knowledge, the database for the two important examination tasks, namely color recognition and trait recognition, was absent before our work. The color and trait of the macroscopic feces samples in our database were carefully labeled by professional doctors, so the database has a high research and medical value. In addition, the results on it are reliable. For the color recognition, the images are labeled into ‘Black’, ‘Red’ and ‘Normal’, the samples are shown in Fig. 2a-c, respectively. For the trait recognition, the images are labeled into ‘Loose’, ‘Watery’ and ‘Normal’, the samples are shown in Fig. 2d-f, respectively.

(2) Data augmentation and over-sampling

Although lots of macroscopic feces images were collected, the database is still class-imbalanced because the probabilities of the occurrence of the feces classes are highly different. The traditional data augmentation methods are unsuitable for this task, because the localization of the tube and the camera imaging conditions are fixed in the controlled environment of our machine. Hence, it is unreasonable to directly arbitrarily rotate or add noise for data augmentation. Generative adversarial network (GAN) is widely used to augment data [39]. There are two main disadvantages for GAN-based methods: 1. A large number of training data are

![Fig. 2 Samples of different colors and traits.](image-url)
typically required; 2. The generated images sometimes are different from the real images. In this paper, we develop a data augmentation strategy to augment the samples of minority classes, i.e. the classes lacking samples. The scarce images are horizontally mirrored at first, the mirrored images have the tube localization and the image quality that conform to the true environment of our machine. Meanwhile, the scarce samples are over-sampled in each batch according to the probability relationship. The augmentation and over-sampling ensure the generalization ability in the true environment.

(3) Light-weight advantage

Considering the computing resource and the hardware cost of the machine, a light-weight structure is demanded. Tewari and Gupta [46] proposed an ultra-lightweight mutual authentication protocol to reduce storage and communication cost. Masud et al. [37] proposed a lightweight and physically secure authentication key for Internet of Medical Things. The sufficient experiments and discussions demonstrate that this protocol is robust and consumes fewer resources. In general, there are some advantages of a light-weight approach. 1. The requirement of computation hardware is low, so it is easy for the hospital to update the model if more additional data are collected; 2. The computational complexity can be low, and accordingly the computing speed can be fast; 3. The cost of the machine can be significantly reduced, so more hospitals can afford our machine.

(4) Multi-task recognition

It is probable that the color and trait have latent relationship that was neglected in the former works. Thus we develop a fast and accurate multi-task light-weight diagnosis network, in which the parameters are shared by the two tasks, namely color recognition and trait recognition. The parameter number of a single dual-task network is generally much smaller than the total parameter number of two single-task networks. In addition, the multi-task network can mine the latent correlation between the trait and color, so the multi-task network yields higher accuracy and faster converge speed than the single-task networks.

(5) Weighting strategy

Adaptive weighting can enhance the discrimination and improve the accuracy [1]. In the multi-task network, the loss function is the weighted sum of the losses of the two tasks. The weights are determined by the learning abilities of the tasks. The accuracy of a single task on the training set is linearly fitted, the slope and the bias of the fitting function are used to measure the learning ability of the task. The higher weight should be given to the relatively difficult task.

The rests of this paper are organized as follows: Section 2 introduces related works. Section 3 specifies our framework in detail. The experiments and discussions are in Section 4. The conclusions are drawn in Section 5.

2 Related works

The traditional examination relies on lots of manual operations that are typically based on some physical or chemical theories. Berhe et al. [7] evaluated the recognition rate of Kato-Katz
for feces specimen examination, which is recommended as an effective diagnosis method by World Health Organization. Bergquist et al. [6] and Charoenstuk et al. [12] compared different popular feces examination methods, including Kato-Katz, formalin-ethyl acetate concentration technique and fecal parasite concentrator kit. These methods achieved satisfactory performance, but they are difficult to spread to clinics due to some stringent requirements, including a large number of the people with professional skills and expensive consumables.

In terms of cost and accuracy, computer-aided diagnosis is considered as the future generation clinical assistant diagnosis technology. Alomari et al. [2] applied circular Hough transform to count abnormal cells. Carvalho et al. [9] combined support vector machine and fuzzy theory to diagnose lung disease. Baldoumas et al. [4] employed natural time analysis to design a conventional three-electrode electrocardiography system. Yang et al. [50] predicted computed tomography (CT) image from magnetic resonance imaging (MRI) by the nonlinear local descriptors, and achieved satisfactory performance. Liu et al. [35] proposed a deep iterative reconstruction estimation strategy with a 3D residual convolutional network structure to reconstruct the low-dose CT images. Cheng [14] developed a two-step grading scheme. Firstly, the testing image was sparsely represented by the set of the training images, then the image was graded by referencing the selected atoms. Nkamgang et al. [38] applied a neuro-fuzzy classifier to detect intestinal parasite and the input was the histogram of the orientation gradient feature of the microscopic images. Caselli et al. [10] designed an approach based on Bayesian to count the number of cells. Wang et al. [48] developed additive least square support vector machine to reduce the infection of missing data in the community health and epidemiological studies. Ginneken et al. [19] presented an active shape model, which used kNN-classifier to describe the shapes of the tissue or organ in lung and brain images.

The aforementioned methods are based on hand-crafted features, so they require the prior knowledge in the related fields. Deep learning models with self-learned structures, such as CNN, do not need the hand-crafted features, which have robust and higher recognition performance. Deep learning models have achieved excellent performances in different computer vision tasks [23, 44, 47, 49], such as denoising [5], detection [15] and segmentation [53]. Accordingly CNN has been important to medical image analysis [20].

Kermany et al. [26] established a deep learning framework to diagnose common treatable blinding retinal diseases. Kumar et al. [28] proposed a novel CNN feature extractor that had been pre-trained from natural images, and then the model was fine-tuned and optimized for the specific medical task. Tajbakhsh et al. [45] used pre-trained CNNs for the classification, detection, and segmentation in four distinct medical image datasets. Effective and reasonable fusion can improve the recognition accuracy [29, 30]. Chang et al. [11] fused the CT image domain information and the corrected images based on a CNN structure to generate hybrid images. Savelli et al. [42] presented a multi-context ensemble CNN to detect small lesions in CT images, different image patches were trained individually, and then these features were combined. With these combined features, CNN can better discover the details in medical images. Qureshi et al. [40] utilized 3D CNN architecture and brain MRI to detect schizophrenia discriminate patients with schizophrenia and normal healthy control subjects. Their approach achieved a high detection accuracy. Huang et al. [24] developed a two-stage scheme based on stationary wavelet transform and residual neural network to denoise the low dose CT images. Eun et al. [18] presented a novel ensemble CNN for nodule detection, and a non-
nodule categorization was proposed to extend the learning capability of the network. Liu et al. [33] developed a novel cascaded deep CNN to complete two challenging tasks simultaneously, namely the automatic segmentation of brainstem gliomas and the prediction of genotype mutation status based on MRI. Liu et al. [34] developed a deep multi-task multi-channel learning framework for brain disease classification and clinical score regression, which used MRI data and demographic information. Jodeiri et al. [25] studied the Mask R-CNN on radiography image segmentation, meanwhile, transfer learning, multi-task learning and data augmentation were used to improve the learning ability of the network. Reena and Ameer [41] proposed a deep learning-based method to localize and recognize the leukocytes in peripheral blood.

Different from the other vision tasks, it is highly difficult to collect the database of medical image due to two main reasons. 1. The number of the patients with a specific disease is small. 2. The images contain much private information, so patients and their family are usually unwilling to share their pathological samples, and it is highly difficult to obtain the approval agreements from the patients, patients’ families and hospitals. Hence, the scarcity of feces sample seriously hinders the research progress of macroscopic feces examination. Thus few studies focus on this field even though it is an important clinical examination.

Hachuel et al. [21] utilized ResNet [22] to recognize feces traits. They collected a large database, the images were labeled into ‘Constipation’, ‘Loose’ and ‘Normal’. However, these images were collected by the patients in the uncontrolled environments, accordingly this method must be designed as a deep structure to overcome the interference in the uncontrolled environment. Yang et al. [51] proposed StoolNet to recognize the color of the feces images. StoolNet as a shallow CNN was carefully designed for the feces samples collected in real hospitals. Inspired by StoolNet, Leng et al. [31] collected a feces database with various traits, and developed a shallow CNN that was suitable for feces trait recognition.

As discussed above, the existing works mainly rely on a uniform distribution data, so the class-imbalanced problem is ignored. In addition, these methods cannot recognize the color and the trait of the feces images simultaneously, so the complementarity between two tasks was not leveraged. Meanwhile, the hand-craft-based classifiers are not robust enough, it is difficult to employ them into the real clinical environments. Besides, the weight selection in CNN-based multi-task learning is empirical, there is not a practical guidance to select appropriate weights for different tasks.

3 Methodology

3.1 Framework

Automatic medical clinical diagnosis is one of the most important research fields in smart city [3, 17, 52]. In order to reduce the burden of the doctors, feces clinical diagnosis is studied in this paper. A multi-task light-weight CNN is developed for the color and trait recognition of macroscopic feces images. A practical data augmentation method and an over-sampling strategy are employed to solve the sample scarcity of minority classes. Meanwhile, the weights for the losses of the tasks are determined according to the task difficulties. The whole framework is shown in Fig. 3.
3.2 Data augmentation and over-sampling strategy

Data augmentation is necessary because the number of pathological samples, especially minority class samples, is very limited. In the typical augmentation methods, the images are rotated or added noise; however, they are unsuitable for our machine. In our machine, the tube localization is fixed, so the rotated images are unreal, such as 90°, 180° and 270° rotation. Hence, the images of the minority classes are mirrored. Figure 4 shows the sample of a mirrored image.

Unfortunately, even after data augmentation, the augmented database is still class-imbalanced because the occurrence probabilities of the feces classes are highly different. Random sampling is widely used in training, but it would degrade convergence and the recognition accuracy. If the database is significantly class-imbalanced, it is highly probable that the training batch does not contain enough samples of minority classes, then the training...
batch forces the network to focus on the majority classes with enough samples. A statistic over-sampling is proposed and employed for training batch selection to alleviate the class-imbalance. \( P_i \) is the probability that the \( i \)-th class is selected, \( P_{i,j} \) is the probability that the \( j \)-th sample of the \( i \)-th class is selected, then

\[
P_i = \sum_{j=1}^{m} P_{i,j}
\]

(1)

where \( m \) is the number of classes, \( m=3 \) in this paper. To ensure class-balance, i.e. each class is selected with the same probability, then

\[
P = \frac{1}{m}
\]

(2)

Thus

\[
P_i = \frac{P}{c_i}
\]

(3)

where \( c_i \) is the number of the samples of the \( i \)-th class. With the over-sampling, the minority classes are selected more frequently, and the majority classes are selected less frequently. Accordingly, the numbers of all classes in the training batch are the same.

In order to help readers to trace our oversampling strategy, the pseudo code is shown as follows:

---

**Algorithm 1: Over-sampling Strategy**

**Input:** \( D_i^j \) is the \( j \)-th sample of the \( i \)-th class, \( m \) classes, \( c_i \) is the number of \( i \)-th class, \( n \) is batch size

**Output:** Training batch \( T \)

\[
P \leftarrow \frac{1}{m}
\]

for \( i \leftarrow 1 \) to \( m \)

\[
P_i \leftarrow \frac{P}{c_i}
\]

end for

for \( i \leftarrow 1 \) to \( n \)

\[
r \leftarrow \text{rand}[0,1]
\]

\[
c \leftarrow \text{ceil}(r/P)
\]

\[
z \leftarrow r - c \times P
\]

\[
j \leftarrow \text{ceil}(z/P_c)
\]

\[
T_i \leftarrow D_i^j
\]

end for

---
3.3 Weight selection strategy

The loss of multi-task network is the weighted sum of the losses of two tasks. The weights are determined according to the learning abilities of the tasks. The higher weight should be given to the relatively difficult task. Generally, the learning ability can be measured by the learning speed and the accuracy. The accuracy increment in training reflects the learning speed, so a mathematical model based on accuracy increment is established for quantitative analysis on the task difficulty.

The accuracies and corresponding epochs in training are used as the ordinate values and abscissa values, respectively, for linear function fitting. The least square method is adopted to calculate the slope \( k \) and the bias \( b \).

\[
k = \frac{\sum xy - \frac{1}{N} \sum x \sum y}{\sum x^2 - \frac{1}{N} (\sum x)^2}
\]

where \( x \) is the epoch, \( y \) is the accuracy.

\[
b = \bar{y} - k \bar{x}
\]

where \( \bar{y} \) is the average accuracy and \( \bar{x} \) is the average epoch value.

If the epoch number of the durative accuracy oscillation is less than \( t \), then the network can be considered as converged. If the accuracies of continuous \( t \) epochs are greater than 95\%, it is considered as similar convergence. The number can be set as different values for different tasks. The main reason is that if the training epoch is set very great, the influence of epoch number for the fitted function is much greater than that of accuracy. Hence, if the network is converged before the epoch number reaches the maximum epoch, \( N \) is the epoch where the network reaches convergence; or else \( N \) is the maximum epoch.

It is easy to understand that different tasks have different difficulty level. If one task is more difficult than the other one, the network should pay more attention to the more difficult one. Hence, a quantitative difficulty assessment is necessary. The fitted linear function reflects the learning speed and the learning ability of the network for a specific task; while the slope and the bias can jointly reflect the learning ability of a specific task. It is obvious that the greater the two values are, the better the learning ability of the network for the task is. The network should focus on the difficult task more than the easy one, so the loss weight of the difficult task should be higher than that of the easy one.

3.4 Multi-task light-weight CNN

Compared with the traditional recognition methods, CNN is typically more robust and accurate. For the traditional methods, it is almost impossible to intuitively know which features are ideal to the specific tasks, especially multi-task task. It is difficult to extract discriminate features that are generally abstract and contain semantic information. CNN is free from hand-crafted design and could ensure robustness. Meanwhile, the nonlinear multi-layer structure can effectively learn extracting strong semantic information, so the recognition performance of CNN is impressive. Hence, CNN is selected as the basic structure in our method.
A multi-task light-weight CNN is developed to recognize macroscopic feces images, the network structure has been shown in Fig. 3. The input of the network is resized as 100 × 100, and the original image size is 480 × 640. The network is composed of three convolutional layers and a fully connected layer, the numbers of filters on the three convolutional layers are 32, 64 and 128, respectively. The sizes of convolutional filters are all 3 × 3. The input of the fully connected layer is a 256-dimensional vector, the numbers of the outputs for the color recognition and trait recognition are both 3. Batch size is set to 64 and the learning rate is set to 0.0001. In order to conveniently reproduce the structure, the key functions are defined as follows:

ReLU is the activation function:
\[
f(x) = \max(0, x)
\]  
Dropout is used on fully connected layers to avoid the over-fitting problem. Dropout is a regularization approach in the training CNNs by randomly dropping out some nodes during the training of the network [43]. With dropout, the training is done with a set of networks. This method, similar to ensemble learning, is powerful in reducing over-fitting. Dropout forces the neurons to learn the features of the input dependently, so each neuron do not rely on the other neurons. Here the dropout rate is set as 0.5. The outputs of the network are normalized by softmax function:
\[
softmax(y_i) = \frac{e^{y_i}}{\sum_{j=1}^{m} e^{y_j}}
\]

where \(i\) is the class index, \(y_j\) is the probability of the \(i\)-th class, \(m\) is the number of classes.

### Table 1 The slopes and biases of the fitted linear functions for two tasks

| Task            | Slope       | Bias       | Regressed value |
|-----------------|-------------|------------|-----------------|
| Color recognition | 1.41 × 10^{-4} | 0.9179    | 0.9602          |
| Trait recognition | 2.01 × 10^{-4} | 0.9237    | 0.9840          |

Fig. 5 Fitting results of two single tasks. a Color. b Trait

![Graphs](image-url)
4 Experiments and discussions

The experiment configuration includes Intel Core i5-7400 U CPU, 8GB internal storage, NVIDIA GeForce GTX 1050 GPU. All codes were written in Python, and the deep learning framework was TensorFlow.

It is highly difficult to collect medical feces images and find professional doctors who can label the samples both in color and trait. We collected a feces image database containing 435 images, in which all samples were labeled by professional doctors in color and trait. 75% images are used as the training samples; while the other images are used as testing samples. Our method is trained on the training data, and then our approach is tested on the testing data to assess the accuracy and efficiency.

It is important to measure the learning abilities of the model for different single tasks, and the weights for different tasks are roughly selected by their learning abilities. As the discussion in Section 3.3, $t=5$ and the maximum epoch is set to 300. Figure 5 shows the fitting results of two single tasks, namely color and trait recognition, on the non-augmented database. Color recognition is more difficult than trait recognition according to the slopes and the biases in Table 1. However, the level of unbalanced data probably affect the results. Unbalanced data are high unfriendly to minority classes. In the following experiments, we explore the influence of unbalanced data.

![Fig. 6](image_url) The fitting samples for the single recognition task. a Color recognition; b Trait recognition

### Table 2  The slopes and biases of the fitted linear function with different weights

| Weight (Color) | Slope (Color)   | Bias (Color) | Slope (Trait) | Bias (Trait) |
|----------------|-----------------|--------------|---------------|--------------|
| 0              | $6.17 \times 10^{-5}$ | 0.2817       | 0.0017        | 0.9257       |
| 0.1            | 0.0024          | 0.7710       | 0.0034        | 0.8933       |
| 0.2            | 0.0055          | 0.6736       | 0.0050        | 0.8527       |
| 0.3            | 0.0095          | 0.6700       | 0.0043        | 0.8992       |
| 0.4            | 0.0053          | 0.7308       | 0.0029        | 0.8859       |
| 0.5            | 0.0141          | 0.6453       | 0.0030        | 0.9030       |
| 0.6            | 0.0023          | 0.8442       | 0.0048        | 0.8795       |
| 0.7            | 0.0028          | 0.8441       | 0.0037        | 0.8636       |
| 0.8            | 0.0014          | 0.8844       | 0.0072        | 0.8164       |
| 0.9            | 0.0039          | 0.8236       | 0.0043        | 0.8446       |
| 1.0            | 0.0041          | 0.8444       | $-5.87 \times 10^{-5}$ | 0.2462       |
Figure 6 shows the fitting results of two tasks with data augmentation, which demonstrates that our data augmentation improves the convergence, including accelerating convergence speed and reducing the convergence difficulty. The slopes and the biases with different weights are shown in Table 2. The difficulty levels of the two tasks in the multi-task lightweight CNN are similar, so their weights should also be similar.

If the class-imbalance problem is not solved, it is difficult to extract the discriminate features of minority classes. Data augmentation can relieve class-imbalance problem. Benefit from the data augmentation strategy, the single-task networks can converge stably and fast. However, the single task cannot converge without data augmentation strategy, which can be observed in Fig. 5.

In order to intuitively compare the regressed values, Fig. 7 shows the observed values of the fitted function at 300-th epoch with different weights. In general, the accuracies of the multi-task lightweight CNN for two tasks are satisfactory when the weights for two tasks are similar. Besides, it also indicates that the two clinical examinations are highly related, because the learning performance of the network is very similar. The single-task learning cannot converge very fast, while the multi-task learning can converge much faster. Because the learned knowledge from the multi-task training simultaneously improves the accuracies of both two tasks.

Figure 6 The fitting results of different weights

| Task                | Multi-task training With augmentation | Single-task training Without augmentation | Existing methods |
|---------------------|---------------------------------------|------------------------------------------|------------------|
| Color recognition   | 100%                                   | 93.6%                                    | 89.9%            | 96.8% [51]     |
| Trait recognition   | 94.9%                                  | 90.7%                                    | 94.9%            | 93.8% [31]     |
The accuracies on testing set are shown in Table 3. The multi-task strategy can significantly improve the performance, especially in the clinical diagnose tasks, because there is a latent relationship between different clinical examinations, such as feces color examination and feces trait examination. The information of different tasks can mutually complement each other and improve the recognition performance. Because the network is shareable for different tasks, the neurons jointly extract more robust and general features.

Figure 8 reveals the relationship between the network depth and the accuracy of multi-task network with augmentation. In this experiment, all the hyperparameters are fixed. The results demonstrate that “deep is not always good”. Even two tasks have similar difficulty levels, their suitable structures are possible various. Besides, even a task is fixed, its difficulty level in different network structures are also possible various. For the color recognition task, a shallow network is more suitable; while, for the trait recognition task, a deep network is more appropriate. Hence, for multi-task recognition, especially there are some latent relationships between the two tasks, a difficulty level assessment is necessary.

In addition, the efficiencies are compared in Table 4. The running time is strongly related with the hardware. Our running time is satisfactory and there is an obvious space to improve the performance if higher performance hardware is used. Our method is a little slower than the existing state-of-the-art feces diagnosis methods. However, both [31, 51] are for single task, our method can diagnose the trait and color simultaneously and has higher accuracies, so our running time is much less than the sum running time of the two methods.

![Figure 8](image)

**Fig. 8** Relationship between the network depth and the accuracy

|         | [51]   | [31]   | Ours    |
|---------|--------|--------|---------|
| Running time | 0.1723s | 0.1805s | 0.1960s |
5 Conclusions

Feces recognition is an important clinical examination. A multi-task light-weight automatic approach is proposed to recognize abnormal color and trait of the macroscopic feces images. A practical data augmentation method and an over-sampling training strategy are developed to overcome the scarcity of the pathological samples. At the same time, the threshold strategy is proposed based on the linear functions. The multi-task light-weight structure requires much less parameter number than the total parameter number of several single-task structures. Besides, the recognition accuracies of two tasks are both improved by the multi-task network. Our method can be easily modified to embed into the other tasks, such as biometric recognition, clinical diagnosis and image classification. If the two tasks are related, i.e., there is a latent relationship between different tasks, the multi-task network can achieve satisfactory performance. Meanwhile, the proposed data augmentation and oversampling method can also be utilized to other unbalanced-data tasks.

However, our method is designed especially for a controlled environment of a medical clinical machine. For users’ self-detection, the imaging devices and the environments are different and complexity. Hence, in the future works, we would try to develop the method that can be used for users’ self-detection.

Acknowledgements This research was funded by the National Natural Science Foundation of China (61866028, 61866025, 61663031, 62162045), Key Program Project of Research and Development (Jiangxi Provincial Department of Science and Technology) (20192BBE50073).

References

1. Abdulnabi AH, Wang G, Lu J et al (2015) Multi-task CNN model for attribute prediction. IEEE Trans Multimedia 17(11):1949–1959. https://doi.org/10.1109/TMM.2015.2477680
2. Alomari YM, Abdullah S, Huda SN et al (2014) Automatic detection and quantification of WBCs and RBCs using iterative structured circle detection algorithm. Comput Math Methods Med 2014:979302. https://doi.org/10.1155/2014/979302
3. Alsmirat MA, Al-Alem F, Al-Ayyoub M et al (2019) Impact of digital fingerprint image quality on the fingerprint recognition accuracy. Multimed Tools Appl 28:3649–3688. https://doi.org/10.1007/s11042-017-5537-5
4. Baldoumas G, Peschos D, Tatis G, Chronopoulos SK, Christofilakis V, Kostarakis P, Varotsos P, Sarlis NV, Skordas ES, Bechlioulis A, Michalis LK, Naka KK (2019) A prototype photoplethysmography electronic device that distinguishes congestive heart failure from healthy individuals by applying natural time analysis. Electronics 8(11):1288. https://doi.org/10.3390/electronics8111288
5. Bao J, Luo L, Zhang Y, Yang K, Peng C, Peng J, Li R (2021) Half quadratic splitting method combined with convolution neural network for blind image deblurring. Multimed Tools Appl 80:3489–3504. https://doi.org/10.1007/s11042-020-09821-6
6. Bergquist R, Johansen MV, Utzinger J (2009) Diagnostic dilemmas in helminthology: what tools to use and when. Trends Parasitol 25(4):151–156. https://doi.org/10.1016/j.pt.2009.01.004
7. Berhe N, Medhin G, Erko B, Smith T, Gedamu S, Bereded D, Moore R, Habte E, Redda A, Gebre-Michael T, Gundersen SG (2004) Variations in helminth faecal egg counts in Kato-Katz thick smears and their implications in assessing infection status with Schistosoma mansoni. Acta Trop 92(3):205–212. https://doi.org/10.1016/j.actatropica.2004.06.011
8. Black CJ, Ford AC (2020) Global burden of irritable bowel syndrome: trends, predictions and risk factors. Nat Rev Gastroenterol Hepatol 16:473–486. https://doi.org/10.1038/s41575-020-0286-8
9. Carvalho AOF, Silva AC, Paiva AC et al (2017) Lung-nodule classification based on computed tomography using taxonomic diversity indexes and an SVM. J Signal Process Syst 87(2):179–196. https://doi.org/10.1007/s11265-016-1134-5
34. Liu J, Chen F, Pan C, Zhu M, Zhang X, Zhang L, Liao H (2018) A cascaded deep convolutional neural network for joint segmentation and genotype prediction of brainstem gliomas. IEEE Trans Biomed Eng 65(9):1943–1952. https://doi.org/10.1109/TBME.2018.2845706

35. Liu J, Zhang Y, Zhao Q, Lv T, Wu W, Cai N, Quan G, Yang W, Chen Y, Luo L, Shu H, Coatrieux JL (2019) Deep iterative reconstruction estimation (DIRE): approximate iterative reconstruction estimation for low dose CT imaging. Phys Med Biol 64(13):135007. https://doi.org/10.1088/1361-6560/ab18db

36. Mao R, Qu Y, He JS, Tan JY, Li XH, Liang J, Shen J, Zhu LR, Chen Y, Iacucci M, Ng SC, Ghosh S, Chen MH (2020) Manifestations and prognosis of gastrointestinal and liver involvement in patients with COVID-19: a systematic review and meta-analysis. The Lancet Gastroen Hepatol 5(7):667–678. https://doi.org/10.1016/S2468-1253(20)30126-6

37. Masud M, Gaba GS, Alqahtani S et al (2020) A lightweight and robust secure key establishment for internet of medical things in COVID-19 patients care. IEEE Internet Things. https://doi.org/10.1109/JIOT.2020.3047662

38. Nkamgang OT, Tchiotsop D, Fotsin HB, Talla PK, Louis Dorr V, Wolf D (2019) Automatic the clinical stools exam using image processing integrated in an expert system. Inform Med Unlocked 15:100165. https://doi.org/10.1016/j.imu.2019.100165

39. Pollastri F, Bolesi F, Paredes R, Grana C (2020) Augmenting data with GANs to segment melanoma skin lesions. Multimed Tools Appl 79:15575–15592. https://doi.org/10.1007/s11042-019-7717-y

40. Qureshi MNI, Oh J, Lee B (2017) 3D-CNN based discrimination of schizophrenia using resting-state fMRI. Artif Intell Med 98:90–107. https://doi.org/10.1016/j.artmed.2019.04.003

41. Reena MR, Ameer PM (2020) Localization and recognition of leukocytes in peripheral blood: a deep learning approach. Comput Biol Med 126:104034. https://doi.org/10.1016/j.compbiomed.2020.104034

42. Savelli B, Bria A, Molinara M, Marrocco C, Tortorella F (2020) A multi-context CNN ensemble for small lesion detection. Artif Intell Med 103:101749. https://doi.org/10.1016/j.artmed.2019.101749

43. Srivastava N, Hinton G, Krizhevsky A et al (2014) Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res 15(1):1929–1958

44. Su A, He X, Zhao X (2021) JPEG steganalysis based on ResNeXt with gauss partial derivative filters. Multimed Tools Appl 80:3349–3366. https://doi.org/10.1007/s11042-020-09350-2

45. Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, Liang J (2016) Convolutional neural networks for medical image analysis: full training or fine tuning. IEEE Trans Med Imaging 35(5):1299–1312. https://doi.org/10.1109/TMI.2016.2535302

46. Tewari A, Gupta BB (2017) Cryptanalysis of novel ultra-lightweight mutual authentication protocol for IoT devices using RFID tags. J Supercomput 73:1085–1102. https://doi.org/10.1007/s11227-016-1849-x

47. Tian C, Xu Y, Zuo W (2020) Image denoising using deep CNN with batch renormalization. Neural Netw 121:461–473. https://doi.org/10.1016/j.neunet.2019.08.022

48. Wang G, Deng Z, Choi KS (2016) Tracking missing data in community health studies using additive LS-SVM classifier. IEEE J Biomed Health Inform 22(2):579–587. https://doi.org/10.1109/JBHI.2016.2634587

49. Wong WJ, Lai SH (2020) Multi-task CNN for restoring corrupted fingerprint images. Pattern Recognit 101:107203. https://doi.org/10.1016/j.patcog.2020.107203

50. Yang W, Zhong L, Chen Y, Lin L, Lu Z, Liu S, Wu Y, Feng Q, Chen W (2018) Predicting CT images from MRI data through feature matching with learned nonlinear local descriptors. IEEE Trans Med Imaging 37(4):977–987. https://doi.org/10.1109/TMI.2018.2790962

51. Yang Z, Leng L, Kim BG (2019) StoolNet for color classification of stool medical images. Electronics 8(12):1464. https://doi.org/10.3390/electronics8121464

52. Yu C, Li J, Li X, Ren X, Gupta BB (2018) Four-image encryption scheme based on quaternion Fresnel transform, chaos and computer generated hologram. Multimed Tools Appl 77:4585–4608. https://doi.org/10.1007/s11042-017-4637-6

53. Zhang Y, Chu J, Leng L, Miao J (2020) Mask-refined R-CNN: a network for refining object details in instance segmentation. Sensors 20(4):1010. https://doi.org/10.3390/s20041010

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.