Lecture Notes in Electrical Engineering 1043

Wei Wang
Jiasong Mu
Xin Liu
Zhenyu Na Na *Editors*

Artificial Intelligence in China

Proceedings of the 5th International Conference on Artificial Intelligence in China

Springer

Lecture Notes in Electrical Engineering

Volume 1043

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering—quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

China

Jasmine Dou, Editor (jasmine.dou@springer.com)

India, Japan, Rest of Asia

Swati Meherishi, Editorial Director (Swati.Meherishi@springer.com)

Southeast Asia, Australia, New Zealand

Ramesh Nath Premnath, Editor (ramesh.premnath@springernature.com)

USA, Canada

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

**** This series is indexed by EI Compendex and Scopus databases. ****

Wei Wang · Jiasong Mu · Xin Liu · Zhenyu Na Na
Editors

Artificial Intelligence in China

Proceedings of the 5th International Conference on Artificial Intelligence in China

Springer

Editors
Wei Wang
College of Artificial Intelligence
Tianjin Normal University
Tianjin, China

Jiasong Mu
Tianjin Normal University
Tianjin, China

Xin Liu
Dalian University of Technology
Dalian, China

Zhenyu Na Na
School of Information Science
and Technology
Dalian Maritime University
Dalian, China

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-99-7544-0 ISBN 978-981-99-7545-7 (eBook)
https://doi.org/10.1007/978-981-99-7545-7

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Paper in this product is recyclable.

Contents

AgsNet: An Attention-Guided Lightweight Segmentation Network

Minghui Li, Zengmin Xu, Yichuan Zhang, Lingli Wei, Ningjie Zhou, and Yanan Cui

Abstract Urinalysis test strips are commonly used for urine routine examination. However, due to possible defects in the liquid path, such as blockages, droplets may leak during the process of dropping urine samples onto the test strips, which severely affects the results of medical tests. Therefore, we propose the Attention-guided segmentation Network (AgsNet) to address errors in medical test results caused by defects in the liquid path. AgsNet adapts its focus to different areas of the test strip image, effectively extracting a richer and more diverse set of features. The best segmentation result is obtained with the AgsNet achieving a mean Intersection over Union (mIoU) score of 71.8 and mean Average Precision (mAP) scores of 84.49, respectively. These results underscore AgsNet's potential in significantly reducing the impact of liquid pathway defects on the reliability of urinalysis test outcomes.

Keywords MobilenetV2 · DeepLabv3+ · Attention mechanism · Bayesian

M. Li · Z. Xu
Guangxi Colleges and Universities Key Laboratory of Data Analysis and Computation, School of Mathematics and Computing Science, Guilin University of Electronic Technology, Guilin 541004, China

Z. Xu (✉)
Center for Applied Mathematics of Guangxi (GUET), Guilin 541002, China
e-mail: xzm@guet.edu.cn

L. Wei · N. Zhou · Y. Cui (✉)
School of Artificial Intelligence, Guilin University of Electronic Technology, Guilin 541004, China
e-mail: YananCui@guet.edu.cn

Y. Zhang
School of Computer Science and Information Security, Guilin University of Electronic Technology, Guilin 541004, China

W. Wang et al. (eds.), *Artificial Intelligence in China*, Lecture Notes in Electrical Engineering 1043, https://doi.org/10.1007/978-981-99-7545-7_5

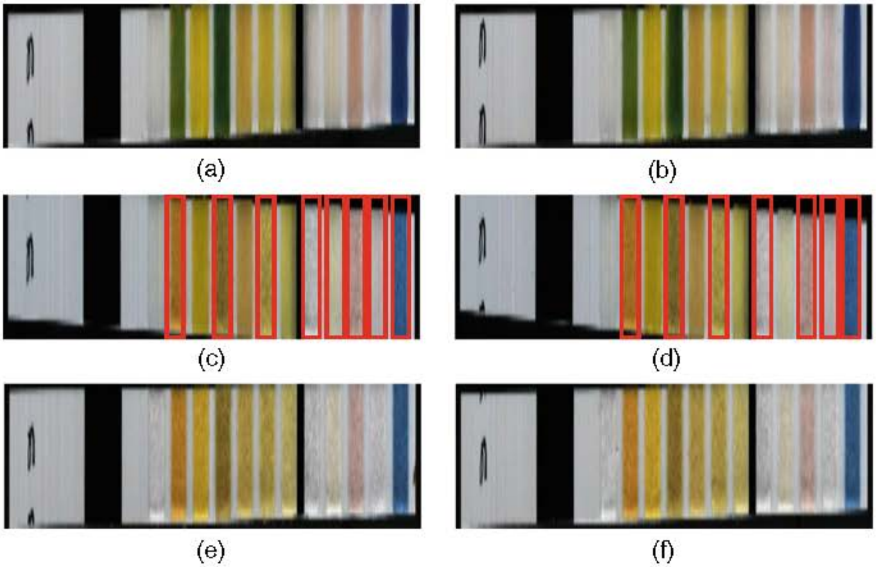

Fig. 1 **(a)(b)** Represent examples of fully-dripped cases, **(c)(d)** Depict partly-dripped cases, and **(e)(f)** illustrate un-dripped cases.

1 Introduction

Urine testing is an important component of contemporary medical clinical diagnosis [1], and the automatic urine analyzer is a machine designed to streamline this process. The machine is composed of multiple systems [2], including a liquid circuit system, a strip selection system, a detection system, an electronic system, and a software system. The liquid circuit system is responsible for dispensing the urine onto the test strip and consists of various pipelines and pump valves. However, due to the weak acidity of urine and the potential for long-term use of the machine, there may be issues with pipeline aging, residual urine components, or other phenomena that can lead to obstructions or poor airtightness in the liquid circuit system, may cause anomalies of urine dripping (Fig. 1).

Differentiating between urine samples containing droplets and those without droplets solely based on visual inspection is a challenging task. Moreover, the limitations of the urine analyzer may result in the possibility of missing droplets. The pixel-level classification capabilities of semantic segmentation models to correctly identify color blocks with missing droplets.

Similar to our work, Fu et al. proposed the DANet [3], which modeled the semantic interdependencies in both spatial and channel dimensions separately, but the use of the model required consideration of computational and memory resource limitations. Chen Li et al. introduced a nested attention mechanism in Unet++ [4], which achieved good segmentation results on the dataset, but also increased the computational complexity. In contrast, our modified model retains the attention mechanism while simplifying the model. Ling et al. [5] introduced a combination weighted cross-entropy loss to further improve the segmentation and recall rates for small defects. We also employ the same method to further improve the performance of our model. Hu et al. [6] proposed a real-time semantic segmentation method based on Joint Pyramid Attention Network (JPANet). However, they still faced challenges such as the computational burden of the segmentation network and missing spatial information in high-level features. We find that incorporating channel attention and performing an eight-fold upsampling approach can partially address the issue of missing spatial information in high-level features, allowing the network to better distinguish details and boundaries.

To the end, we propose AgsNet to extract high-dimensional information and segment color blocks with missing droplets. The presented method uses Deeplabv3+ [7] as the backbone network, and employs the relatively lightweight MobileNetv2 [8] as the encoder. Additionally, two attention mechanism modules are incorporated into the network architecture, which can enhance the accuracy of the model and reduce the number of parameters, making it deployable on lightweight mobile devices.

2 Materials and Method

2.1 Research Subjects

This experiment uses multiple urine test strips with 8–11 different colored blocks, each representing a unique indicator. The test strip images are created through stretching transformations, and the strips in (a)(b) have already been dripped on, resulting in a smoother surface. The red box in (c)(d) highlights a color block that has not been dripped on, resulting in a rougher surface. The strips in (e)(f) show test strips without any liquid dripped on them, and there are slight variations between the strips due to differences in the manufacturing process from various manufacturers.

The experimental data consist of 395 patient urine dipstick images, with size of 880 * 30. These images are randomly partitioned into training and testing sets, with 316 and 79 images respectively.

2.2 Proposed Model

Due to practical limitations, we could only use the low-computational-power MobileNetV2 [9] as the main feature extraction network. MobileNetV2 is an upgraded version of MobileNet, which has a very important feature of using Inverted Residual Blocks. The entire MobileNetV2 consists of Inverted Residual Blocks. The improved network includes both an encoder and a decoder. To further extract its most

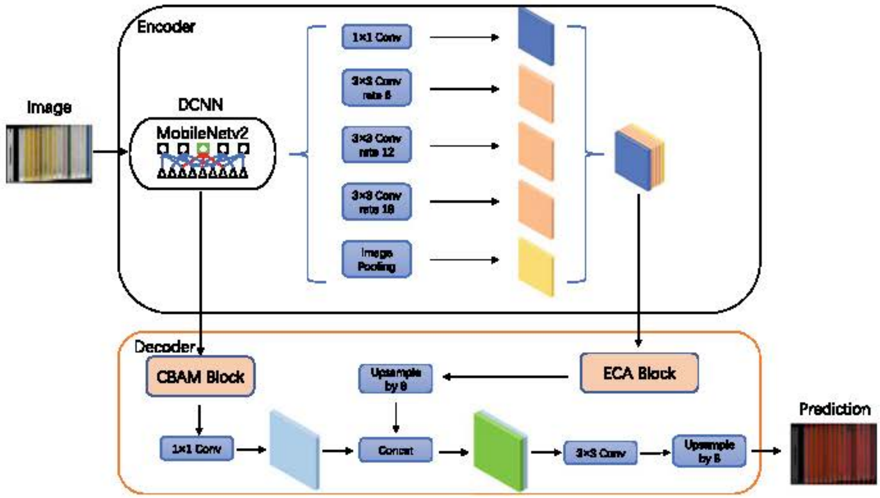

Fig. 2 Based on the Deeplabv3+ model, we have added spatial attention mechanisms and channel attention mechanisms to the decoding part

useful features and enhance the model's understanding of spatial and channel information, we input the extracted shallow features into a CBAM module. This increases the model's perceptual ability, reduces the amount of unrelated information it processes, and improves its performance and efficiency with almost the same number of parameters. Using 1 * 1 convolutions to connect and fuse cross-channel features effectively reduces the number of parameters. During feature extraction, we preserve most of the features between the channel and spatial dimensions, and then fuse them with the deeply extracted features after a 1 * 1 convolution. This strategy allows the model to capture richer and more informative features. The schematic diagram of the entire framework is shown in Fig. 2. We proposed a new formula for the loss function:

$$L = L_{CE} + \lambda_c L_c + \lambda_s L_s \tag{1}$$

whereas L_{CE} denotes the cross-entropy loss function, the weight coefficients for the channel attention loss and spatial attention loss are represented by λ_c and λ_s, respectively.

We add a channel attention mechanism after five stacked features to improve the quality of feature extraction by highlighting important patterns and relationships among channels. This enables our model to capture details and retain important information, enhancing its interpretability and generalization ability. In addition, using an eight-fold upsampling prevent overfitting and improves the model's performance by extracting deeper and more diverse features. Our shallow network feature extraction strategy preserves the most valuable information and effectively integrates it with deep features. These contributions have enabled our model to demonstrate outstanding performance in image segmentation tasks.

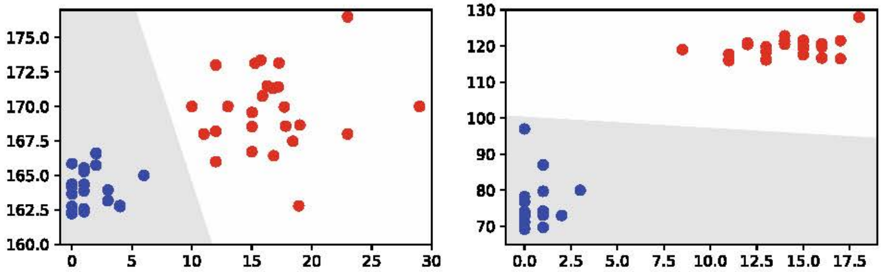

Fig. 3 The classification results of the Bayesian network on two examples, where the blue represents the features before droplet application, and the red represents the features after droplet application

2.3 Bayesian Classifier

We demonstrate in Fig. 3 the application of non-maximum suppression to images using a modified Bayesian network. The process involves calculating the Features from accelerated segment test (Fast) [15] score from the detected feature points, followed by the classification of the resulting features.

We also conducted comparative work on machine learning because the types of urine test strips may vary. We first use the Hash algorithm to identify the correct model of the urine test strip. The similarity between two models is measured by the Hamming distance and the standard library. The smaller the Hamming distance, the higher the similarity. A Hamming distance of 0 means that the two images are identical. We extract the noise of the unused test strip through the Fast. Here is our proposed improved Bayesian classifier:

$$c = \arg\max_{c_i} \frac{P(x, C(x, y)|c_i)}{P(x|c_i)} P(c_i) \tag{2}$$

whereas $P(x, C(x, y)|c_i)$ can be computed based on the sample characteristics x, category c_i, and considering y as the label, we are able to obtain favorable outcomes by utilizing Bayesian networks for classification. However, given the considerable quantity of test strip models, we have currently set aside this approach.

2.4 Attention Mechanism

The introduced attention module enhances performance while avoiding dimensionality reduction by proposing a strategy. One-dimensional convolution is used to reduce parameter computation, and an efficient channel attention block is added. Spatial attention mechanisms are incorporated into SE and ECA blocks, achieving dual-channel attention [10]. The details of the module are shown in Fig. 4. They also use convolutional block attention modules to reduce parameter size [11].

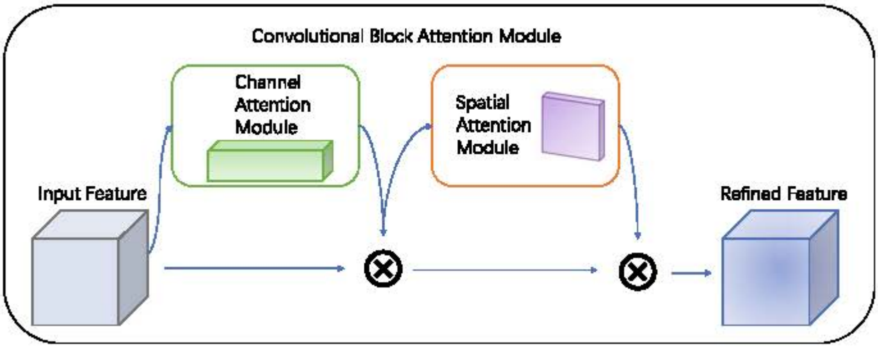

Fig. 4 Attention mechanism

3 Experiment

3.1 Data Preprocessing and Augmentation

By scaling the image to a size of 960 * 256, and then changing the length and width of the image at random, the data enhancement technology includes random rotation, ranging from 1 to 360°, horizontal and vertical flipping used in this study, the image is transferred to HSV color. The first row displays the original images, while the second row shows the detection results. The red bars displayed on the detection results indicate the segmented areas where droplets are missed (Fig. 5).

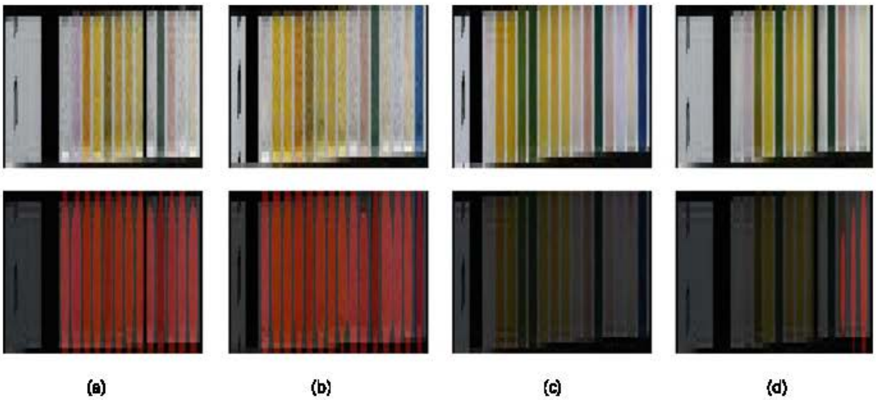

Fig. 5 The red labels represent the segmentation results before the droplet is applied

Table 1 Quantitative evaluation results on the test set. We compare five models and use two evaluation indicators, mIoU and mAP, to analyze the results

Structure	Backbone	mPrecision	Recall	mIoU	mAP	Size (MB)
UNet	VGG	73.29	60.61	48.98	60.61	95.0
	Resnet50	67.07	71.69	59.03	71.69	167.9
PSPNet	MobileNetV2	58.65	60.92	38.46	60.92	**9.3**
	Resnet50	54.77	55.99	34.37	55.99	178.5
HRNet	W32	69.44	73.93	54.51	73.93	113.3
	W18	67.82	72.48	51.31	72.48	37.3
AgsNet	MobileNetV2	**82.02**	**84.73**	**70.43**	**82.82**	22.4

We compared five models and used two evaluation indicators, mIoU and mAP, to analyze the results

3.2 Implementation Details

The urine test strip used in this experiment contains 8–11 color blocks, with each block representing a different index. The testing process requires an automated urine analyzer to dispense droplets sequentially from the first color block, following the urine system.

The network utilizes Adam [12] as the optimizer, with the model's learning rate set at 5e−4. To prevent overfitting, weight decay is employed. The momentum parameter [13] used internally by the optimizer is set to 1−4 e. We apply cosine annealing to reduce the learning rate [14]. In this experiment, after multiple comparisons, we have chosen to set the number of epochs at 500.

Experimental setup: The experiments are conducted on a machine equipped with an Intel(R) E5-2650 v4 @ 2.20 GHz processor and four NVIDIA TITAN Xp GPUs, each with 12 GB of memory. All networks are implemented using PyTorch.

3.3 Results

The performance of the model is evaluated in this study using mIoU and mAP as metrics. For comparison purposes, we also conducted training and evaluation of UNet, PSPNet, HRNet on the dataset using the same strategy. We also try to change the different backbone network in order to multi-level contrast, in the end we find that the added attention mechanism, reduction of model parameters, only 22.4 MB in size (Table 1).

The table shows quantitative results for the models evaluated. The Unet model achieves a mIoU score of 59.03 and mAP scores of 71.69. The PSPNet model achieves a lower mIoU score of 38.46, and its model parameters are only 9.3 MB, its performance is not satisfactory. The HRNet model achieves a mIoU score of 54.51, and mAP scores of 73.93, but its model parameters are larger at 113.3 MB. Nonetheless, compare to Proposed model, all of these models have significant deficiencies.

3.4 Conclusion

The proposed model is an improved semantic segmentation model that incorporates an attention mechanism and achieves lightweight model parameters by segmenting samples with and without drops. This innovation makes it possible to deploy the model on mobile devices, greatly facilitating healthcare workers. The test results demonstrate that the improved model achieves superior performance compared to the comparative method.

Acknowledgements This research was supported by the National Natural Science Foundation of China (61862015), the Science and Technology Project of Guangxi (AD21220114, AD20159035). Additionally, this project is funded by the National Student Innovation and Entrepreneurship Training Programme (S202210595302, 202310595053).

References

1. Goldstein DE, Little RR, Lorenz RA, Malone JI, Nathan D, Peterson CM, Sacks DB (2004) Tests of glycemia in diabetes. Diab Care 27(7):1761–1773
2. Khejonnit V, Pratumvinit B, Reesukumal K, Meepanya S, Pattanavin C, Wongkrajang P (2015) Optimal criteria for microscopic review of urinalysis following use of automated urine analyzer. Clinica Chimica Acta 439:1–4
3. Fu J, Liu J, Tian H, Li Y, Bao Y, Fang Z, Lu H (2019) Dual attention network for scene segmentation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 3146–3154
4. Li C, Tan Y, Chen W, Luo X, Gao Y, Jia X, Wang Z (2020) Attention unet++: a nested attention-aware u-net for liver ct image segmentation. In: IEEE international conference on image processing (ICIP). IEEE, pp 345–349
5. Ling Z, Zhang A, Ma D, Shi Y, Wen H (2022) Deep Siamese semantic segmentation network for PCB welding defect detection. IEEE Trans Instrum Meas 71:1–11
6. Hu X, Jing L, Sehar U (2022) Joint pyramid attention network for real-time semantic segmentation of urban scenes. Appl Intell 52(1):580–594
7. Chen LC, Zhu Y, Papandreou G, Schroff F, Adam H (2018) Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the European conference on computer vision (ECCV), pp 801–818
8. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L-C (2018) Mobilenetv2: inverted residuals and linear bottlenecks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4510–4520
9. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778
10. Zhou T, Canu S, Ruan S (2021) Automatic covid-19 ct segmentation using u-net integrated spatial and channel attention mechanism. Int J Imag Syst Technol 31(1):16–27
11. Chen L-C, Papandreou G, Schroff F, Adam H (2017) Rethinking atrous convolution for semantic image segmentation. *arXiv preprint* arXiv:1706.05587

12. Kingma DP, Ba J (2014) Adam: a method for stochastic optimization. *arXiv preprint* arXiv:1412.6980
13. He K, Fan H, Wu Y, Xie S, Girshick R (2020) Momentum contrast for unsupervised visual representation learning. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 9729–9738
14. Loshchilov I, Hutter F (2016) Sgdr: stochastic gradient descent with warm restarts. *arXiv preprint* arXiv:1608.03983
15. Viswanathan DG (2009) Features from accelerated segment test (fast). In: Proceedings of the 10th workshop on image analysis for multimedia interactive services, London, UK, pp 6–8