Lecture Notes in Electrical Engineering 1163

Peng You
Shuaiqi Liu
Jun Wang *Editors*

Proceedings of International Conference on Image, Vision and Intelligent Systems 2023 (ICIVIS 2023)

ICIVIS

Springer

Lecture Notes in Electrical Engineering 1163

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. ****

Peng You · Shuaiqi Liu · Jun Wang
Editors

Proceedings of International Conference on Image, Vision and Intelligent Systems 2023 (ICIVIS 2023)

Springer

Editors
Peng You
National University of Defense Technology
Changsha, Hunan, China

Shuaiqi Liu
Hebei University
Baoding, Hebei, China

Jun Wang
Hebei University
Baoding, Hebei, China

ISSN 1876-1100 ISSN 1876-1119 (electronic)
Lecture Notes in Electrical Engineering
ISBN 978-981-97-0854-3 ISBN 978-981-97-0855-0 (eBook)
https://doi.org/10.1007/978-981-97-0855-0

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

Preface

ICIVIS 2023

It is with great pleasure that we introduce "The 3rd International Conference on Image, Vision, and Intelligent Systems," short for ICIVIS 2023, which convened from August 16 to 18, 2023, in the picturesque city of Baoding, China.

The rapid advancement of technology in these domains has reshaped the way we interact with the world around us. This conference provided a fertile ground for scholars, researchers, and professionals from across the globe to converge and exchange their profound insights, groundbreaking research, and visionary ideas. The diversity of topics and methodologies presented truly exemplified the multidisciplinary nature of this field.

We extend our deepest gratitude to the dedicated members of the organizing committee whose meticulous planning and tireless efforts were instrumental in making this conference a resounding success. More importantly, ICIVIS 2023 won't had a success without our sponsor, Hebei University and co-sponsor, University of Jinan. Also, we thank Prof. Xufeng Yao, (Northeastern University, China), Assoc. Prof. Chen Li (Northeastern University, China), and Dr. Jun Wang (Hebei University, China) for their long-term support on ICIVIS 2023. Their commitment to excellence ensured that every aspect of this event was flawlessly executed.

We would also like to express our sincere appreciation to our esteemed keynote speakers whose expertise enriched the conference program: Prof. Yao Zhao (Beijing Jiaotong University, China), Prof. Nannan Wang (Xidian University, China), Prof. Yoshihiro Yamanishi (Nagoya University, Japan), Prof. Hongliang Ren (National University of Singapore, Singapore), and Prof. Xiuling Liu (Hebei University, China). Their enlightening talks not only provided valuable perspectives but also sparked meaningful discussions among participants.

Furthermore, we extend our thanks to all the authors and presenters who contributed their innovative research. Your contributions have not only elevated the quality of this publication but will undoubtedly serve as a valuable resource for scholars and researchers in the field.

Lastly, we want to express our gratitude to all the participants for their active engagement, stimulating discussions, and unwavering enthusiasm. Your presence transformed this conference into a vibrant platform for learning, collaboration, and knowledge exchange.

We hope that this proceedings volume will serve as a catalyst for further exploration and advancements in image processing, computer vision, and intelligent systems. May it inspire future research and foster ongoing collaborations among scholars and practitioners.

Thank you for your invaluable contributions, and we look forward to welcoming you again at the next conference.

Conference Committee of ICIVIS 2023

Introduction

This book presents peer-reviewed articles from the International Conference on Image, Vision and Intelligent Systems (ICIVIS 2023), August 16–18, 2023, held in Baoding, China. It encapsulates a compendium of cutting-edge research and developments at the intersection of image processing, computer vision, and intelligent system, serving for researchers, scholars, professionals, students, and academicians. The proceedings encapsulates a compendium of pioneering research that has led to breakthroughs in image processing, intelligent control and system modeling, and a host of other domains. This multidisciplinary field is dedicated to advancing the understanding and application of algorithms and models that empower machines to interpret visual information with proficiency akin to human perception. As technology marches forward, image, vision, and intelligent systems are poised to play an increasingly pivotal role across numerous industries and facets of our daily lives. This collection of proceedings stands as a testament to the dynamism and potential of this rapidly evolving field.

Contents

Image

High Order Conditional Random Field Based Cervical Cancer Histopathological Image Classification

Shangqi Zhou[1,2], Xintong Li[1,2], Yanyu Fu[1,2], Xinyu Ouyang[1,2], Jinzhu Yang[2(✉)], Marcin Grzegorzek[3,4], and Chen Li[1,2]

[1] Microscopic Image and Medical Image Analysis Group, College of Medicine and Biological Information Engineering, Northeastern University, Shenyang, China

[2] Key Laboratory of Intelligent Computing in Medical Image, Ministry of Education, Northeastern University, Shenyang, China
yangjinzhu@cse.neu.edu.cn

[3] Institute of Medical Informatics, University of Luebeck, Luebeck, Germany

[4] Department of Knowledge Engineering, University of Economics in Katowice, Katowice, Poland

Abstract. Cervical cancer ranks as the fourth most frequently diagnosed cancer. This disease has an extended precancerous stage and can be completely cured and prevented through early detection. Currently, the analysis of histopathological images of cervical cancer relies on manual assessment by pathologists, a subjective and time-consuming process. Moreover, there is limited research on differentiating the severity of cervical cancer in histopathological images. To address these challenges, this study proposes a classification algorithm based on high-order conditional random field for cervical cancer histopathologic images. The algorithm effectively categorizes images into high differentiation, medium differentiation, and low differentiation stages. Three deep learning models, namely VGG-16, Inception-V3, and ResNet-50, are utilized for pre-classification at the image block level. For image-level classification, the Visual Transformer model is employed. Finally, the block-level and image-level classification components are integrated using the conditional random field model, resulting in the high-order conditional random field model. Following training and testing using the dataset, the model achieves an overall accuracy of 75.4%.

Keywords: Cervical cancer · Histopathological images · Deep learning · Conditional random fields · Computer-aided diagnosis

1 Introduction

Cervical cancer ranks as the fourth most commonly diagnosed cancer and the fourth leading cause of cancer-related deaths among women worldwide [1]. Globally, approximately 200,000 women lose their lives to cervical cancer each year, with over 130,000 new cases reported in China annually [2]. The American Cancer Society estimates that in

P. You et al. (Eds.): ICIVIS 2023, LNEE 1163, pp. 3–16, 2024.
https://doi.org/10.1007/978-981-97-0855-0_1

2021, there will be approximately 14,480 individuals diagnosed with cervical cancer and 4,290 deaths from the disease in the United States [3]. However, low-income countries experience an even more serious situation. The progression of cervical cancer typically requires at least ten years, but early detection allows for effective treatment, leading to survival rates approaching 100% [4]. The prognosis of cervical cancer depends on the stage of detection, which is categorized as high, medium, or low-grade differentiation. Therefore, accurately determining the stage of cervical cancer is crucial.

The classification and differentiation of cervical cancer are typically categorized into three stages. Grade I indicates highly differentiated, well-differentiated, and low malignant tumors. Grade II refers to moderately differentiated and moderately malignant tumors. Grade III represents poorly differentiated and highly malignant tumors [5]. Manual screening of cell images is a labor-intensive and time-consuming process, as pathologists must meticulously examine each smear image under a microscope for disease diagnosis. Conducting screenings for large populations requires analyzing a significant number of samples, requiring skilled personnel and consuming substantial time. Despite the existence of advanced automated systems for smear analysis, less developed and middle-income countries with high cervical cancer incidence and mortality rates still lack access to such systems due to the high costs associated with maintenance [6]. As a result, there is a growing emphasis on researching and developing cost-effective and efficient automated screening tools. These tools aim to assist pathologists in analyzing and diagnosing samples more quickly and reliably.

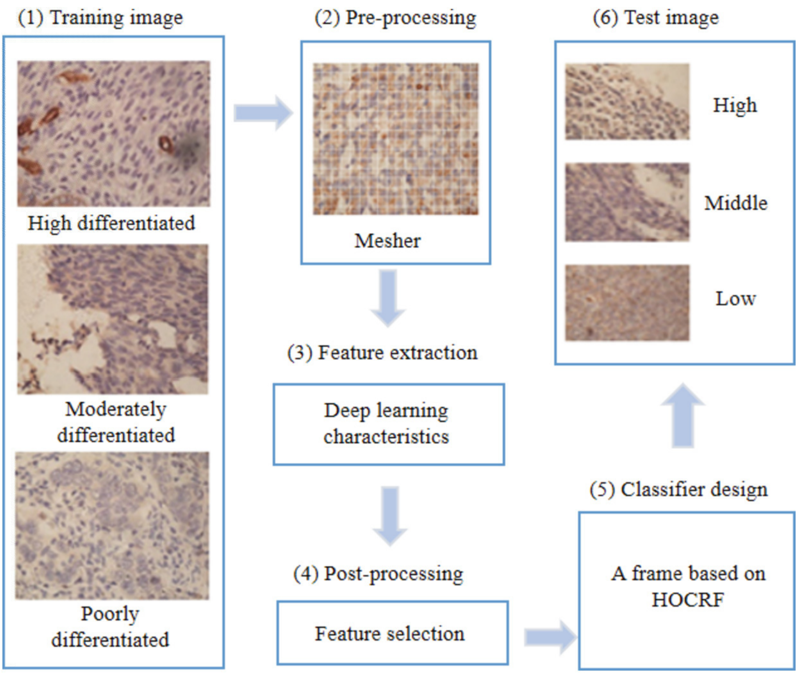

Fig. 1. The workflow of HOCRF.

The research focus of this paper is to use *High-order Conditional Random Field* (HOCRF) model to classify cervical cancer histopathological images into high-differentiation, medium-differentiation and low-differentiation tasks. At present, the application of computer aided diagnosis to classify histopathological images of cervical cancer is generally to classify images with cervical cancer and images without cervical cancer, and there are few classification studies on differentiation level. Therefore, this paper proposes a high order conditional random field framework, as shown in Fig. 1.

2 Related Work

Research status of cervical cancer cytological segmentation methods based on machine learning algorithms is as follows: CNN method based on image blocks is proposed in [7], using selective preprocessing for nuclear segmentation. A method for cell segmentation based on the special resolution of Pap smear images is proposed in [8]. Work of [9] introduces a Progressive Growth U-net model (PGU-net+) to segment the nuclei of cervical cells. [10] uses a new Instant-Relational Network (IR-Net) to segment overlapping cervical cells.

The current research on cytological classification methods for cervical cancer, utilizing machine learning algorithms, can be summarized as follows: In a study by [11], to enhance the classification of cervical cells as normal or abnormal, researchers have introduced the utilization of deep learning and transfer learning techniques. This strategy capitalizes on the capabilities of deep learning models and the transfer of knowledge from pre-existing models to enhance the accuracy of classification. [12] has developed a comprehensive integration model that outperforms previous approaches by comparing and combining the most advanced artificial and non-artificial features. This model has the capability to handle a wide range of image classification problems, including the classification of cervical cells. Another study conducted by focuses on the classification of cervical biopsies, aiming to distinguish between normal and aggressive cancer cases. The researchers utilized three different methods to analyze and classify a total of 475 biopsies, providing insights into the effectiveness of various classification approaches. Overall, these studies contribute to the advancement of cytological classification methods for cervical cancer by incorporating machine learning algorithms. They explore the potential of deep learning, transfer learning, and the integration of advanced features to improve accuracy and provide valuable insights for accurate diagnosis and treatment planning.

Regarding the application of conditional random field (CRF) in the diagnosis and analysis of cervical cancer histopathological images, only Park et al. from the United States have utilized this technique. In their work, they employed a probabilistic graph model based on CRF to analyze histopathological images of cervical cancer. They specifically focused on capturing the spatial relationships and corresponding features among tissues in cervical images. By utilizing CRF, the model was capable of simulating the relationships between neighboring tissues as well as different tissues, thereby demonstrating the spread of cancer tissues to adjacent areas. However, in China, the use of CRF models for detailed classification and categorization of cervical cancer histopathological images has not been explored. Thus far, CRF models in China have primarily been applied to image segmentation and classification in domains other than cervical cancer.

As can be seen from the above, in the direction of cervical cancer histopathological image analysis, it is worth studying to combine with the rapidly developing artificial intelligence methods to assist cervical cancer screening, and the use of conditional random field for AI-assisted diagnosis has great development prospects.

3 Conditional Random Field Module Construction

3.1 Data Preprocessing

The data used in this experiment are the clinical data of Shengjing Hospital, which is affiliated to China Medical University in Shenyang, China. The pretreatment method used in this experiment is meshing. Since the unary potential and binary potential in the high order conditional random field are generated by the image patch level data, we need to obtain the patch level data through the grid method. Due to the limitations of equipment performance, the image of 2560 $\times$ 1920 pixels is first scaled to 1280 $\times$ 960 pixels, which is convenient for computer calculation. With the main message is still retained, the 1280 $\times$ 960 pixel image was cropped into a 100 $\times$ 100 pixel patch-level image, and the parts that were not divisible were cut out. After obtaining the dataset of patch-level images, the next step involves conducting feature extraction.

3.2 Feature Extraction Method

The experiment utilizes various feature extraction methods, including the VGG network, Inception network, Res-Net network, and a relatively new Vision Transformer model. In the VGG-16 network, the default input layer consists of 224 $\times$ 224 $\times$ 3 pixels. The structure begins with two identical convolution layers using a 3 $\times$ 3 pixel convolution kernel, followed by a pooling layer. The process is repeated multiple times, involving two convolution layers and one pooling layer each time. This repetition occurs three times. Following this, three fully connected layers are utilized, resulting in the final output. In total, the VGG-16 network consists of thirteen convolutional layers, three fully connected layers, and five pooling layers. The Inception-V3 network employs a strategy of splitting two-dimensional convolutions into two one-dimensional convolutions, aiming to reduce parameters and mitigate overfitting. In this experiment, the Inception-V3 model extracts a deep learning feature vector with a dimensionality of 1000 dimensions. The ResNet-50 model extracts a feature vector of 1000 $\times$ 1 pixels. The Vision Transformer model applies operations such as convolution and core pooling, transferring the pooling layer from the last layer to the fully connected layer. By considering the number of filters in the fully connected layer, the CNN features of the fully connected layer are obtained in the corresponding number of filter dimensions. In simpler terms, the feature vector is obtained by taking the output of the upper layer of the full convolution as the feature vector.

3.3 Research on Classification Methods based on HOCRF

3.4 Structure of HOCRF

The structure of the high-order conditional random field model proposed in this paper is shown in Fig. 2.

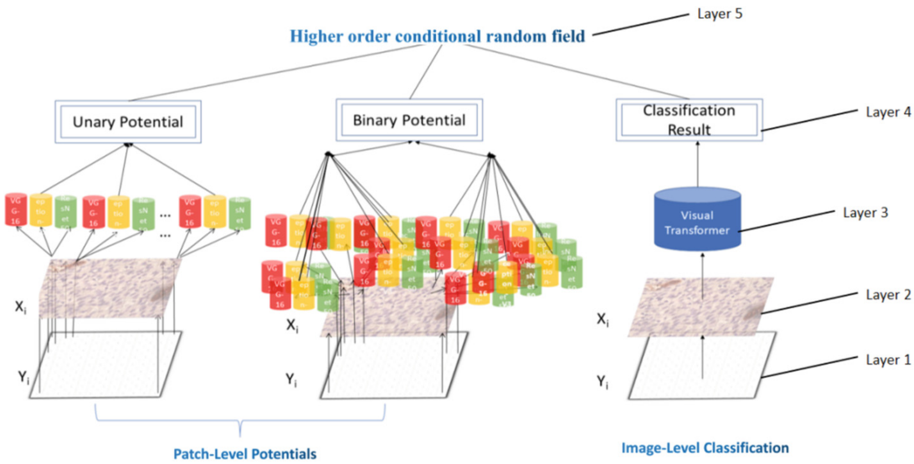

Fig. 2. High order conditional random field framework.

The figure shows the overall structure of the high-order conditional random field model proposed in this experiment. The left side is the block-level image classification, and the right part is the image level classification:

(1) Layer 1: This layer is the label layer of the real label Y_i (i = 1,2,3…) corresponding to each image.
(2) Layer 2: This layer is the visible layer of the original image block Y_i, which corresponds to the label in the first layer.
(3) Layer 3: This layer is the feature extracted from the image X_i in the second layer, and the left patch level image is the deep learning feature extracted from VGG-16, Inception-V3 and ResNet-50. Vision Transform features are extracted from the image-level data on the right.
(4) Layer 4: This layer is used to generate unary potential, binary potential and high-order potential functions with the obtained feature vectors. In the binary potential part, the features of the target image block X_i are obtained by calculating the features of the surrounding image blocks.
(5) Layer 5: the generated unary potential, binary potential and high-order potential are combined in the proposed high-order conditional random field model, and are used as the classifier of cervical cancer histopathological image classification task.

3.4.1 Unary Potential, Binary Potential and Higher Order Potential of HOCRF

Unitary potential: $\varphi_i(x_i; Y)$ of equation

$$P(X|Y) = \frac{1}{Z}\prod\nolimits_{i\in V}\varphi_i(x_i; Y)\prod\nolimits_{(i,j)\epsilon E}\varphi_{i,j}\left(x_i, x_j; Y\right)\prod\nolimits_{i\in V}\varphi_i(x_i, I; Y) \tag{1}$$

(i and j are sequences of real numbers). X and Y are random variables, $P(Y|X)$ is under the condition of a given XY of conditional probability distribution. The random variable Y forms a Markov random field represented by an undirected graph $G = (V, E)$. The probability of the value $c(c \in \mathbb{L})$ of the label y_i is related to it. For the given data X, there is $\varphi_i(y_i; X) \propto p(y_i = c|f_i(X))$, where the image data is represented as the nod-mode feature vector $f_i(X)$, $f_i(X)$ may be determined by all the values of X. The best classification feature is selected according to the pre-classification results, and the best classification result is the monadic potential of the high order conditional random field model.

Binary potential: $\varphi_{i,j}\left(x_i, x_j; Y\right)$ of equation

$$P(X|Y) = \frac{1}{Z}\prod\nolimits_{i\in V}\varphi_i(x_i; Y)\prod\nolimits_{(i,j)\epsilon E}\varphi_{i,j}\left(x_i, x_j; Y\right)\prod\nolimits_{i\in V}\varphi_i(x_i, I; Y) \tag{2}$$

It tells us the probability of $\left(y_i, y_j\right) = \left(c, c'\right)$ for adjacent nodes i and j. The value can be expressed as the following equation $\psi_{ij}\left(y_i, y_j; X\right) = p\left(\left(y_i = c, y_j = c'|f_i(X)f_j(X)\right)\right)$. Figure 3 below shows the structure of a layout of binary potentials. In this experiment, octet neighborhood layout was used, that is, the feature vector of the target block-level image was represented by calculating the sum of feature vectors of eight neighborhood patch level pixel blocks around the target block-level image. The other generation steps are the same as the unary potential.

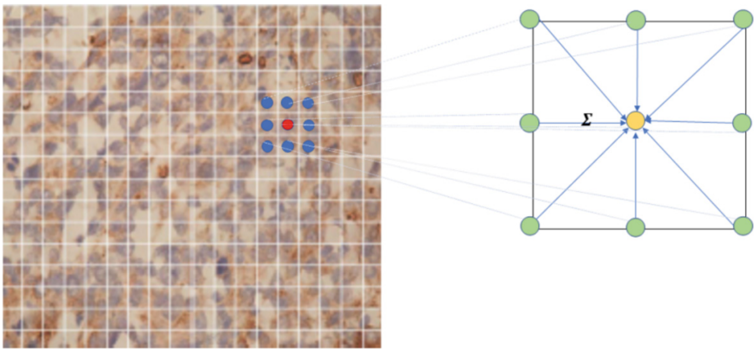

Fig. 3. Layout of the Binary potential.

Higher-order potential: $\varphi_i(x_i, I; Y)$ of equation

$$P(X|Y) = \frac{1}{Z}\prod\nolimits_{i\in V}\varphi_i(x_i; Y)\prod\nolimits_{(i,j)\epsilon E}\varphi_{i,j}\left(x_i, x_j; Y\right)\prod\nolimits_{i\in V}\varphi_i(x_i, I; Y) \tag{3}$$

It represents the spatial relationship between the whole image. Feature vectors and classifiers generated by Vision Transformer model are used as high-order potential functions.

4 Experimental Results and Analysis

4.1 Dataset

In this experiment, the data set was first divided into training set, verification set and test set at the ratio of 1:1:2, and then all the data sets were divided into 100×100 pixel image patches. In order to make full use of the data set, we need to randomly extract from each data set according to the corresponding proportion of high differentiation, medium differentiation and low differentiation. The specific data allocation of image-level data is shown in Table 1.

Table 1. Allocation of experimental data sets.

Data Set		Training Set	Validation Set	Test Set	Total
	high	17	17	34	68
AQP	moderately	15	15	30	60
	poorly	17	17	36	70
	total	49	49	100	198
	high	17	17	35	69
HIF	moderately	16	16	34	66
	poorly	18	18	36	72
	total	51	51	105	207
	high	16	16	32	64
VEGF	moderately	19	19	38	76
	poorly	16	16	33	65
	total	51	51	103	205

4.2 Experimental Results and Analysis of Unary Potential, Binary Potential and High-Order Potential

In the monadic potential section, we use three effective deep learning feature classifiers to describe the monadic potential in this paper. The three deep learning features are VGG-16, Inception-V3 and ResNet-50. In order to compare the classification results of the three classifiers, the average classification index performance assessment will be used next and represented by a line graph. In the following Fig. 4 are the index line graphs of AQP, HIF and VEGF staining methods under different models, in which the

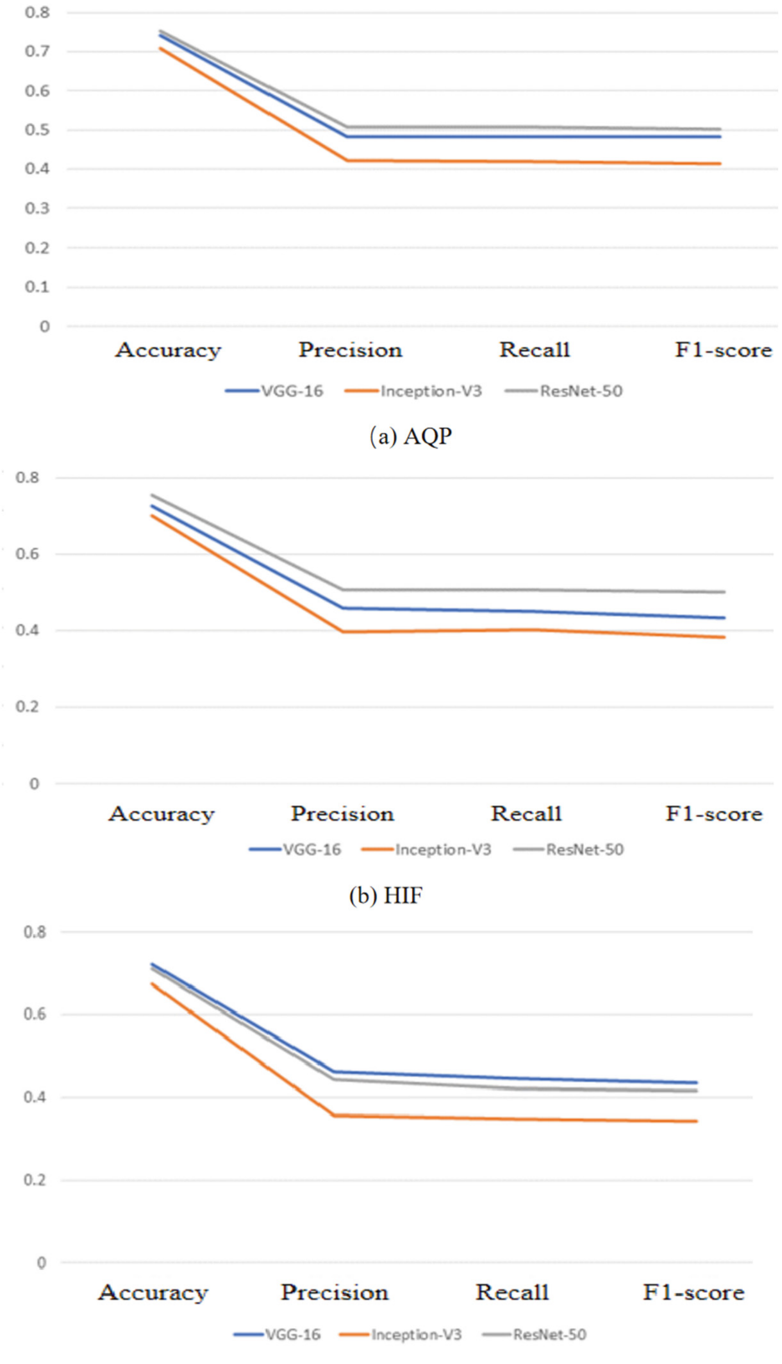

Fig. 4. Line graph analysis of pre-classification results of monadic potential on verification set.

blue line represents the VGG-16 model, the yellow line represents the Inception-V3 model, and the gray curve represents the ResNet-50 model.

The observations from the above figure reveal that the gray curve, representing the ResNet-50 model, is higher than the blue curve and yellow curve in Fig. 4. This indicates that the ResNet-50 model achieves better classification results compared to the VGG-16 and Inception-V3 models in the AQP and HIF staining data sets. On the other hand, in Figure (c), the blue curve is higher than the gray curve and yellow curve, suggesting that the VGG-16 model performs better in the VEGF data set. However, based on the figure, it can be concluded that the classification performance of the ResNet-50 model is nearly equivalent to that of the VGG-16 model. Therefore, considering the overall performance, the ResNet-50 model demonstrates optimal classification performance in the unary potential aspect. Hence, the ResNet-50 model is chosen to construct the unary potential portion.

The experimental procedure for the binary potential follows a similar flow to that of the unary potential, with the exception of incorporating additional layers specific to the binary potential. During the feature extraction step, the features of the target image block are determined by calculating the features of image blocks surrounding the target block-level image. The analysis of the data reveals that the ResNet-50 model exhibits superior classification performance compared to both the Inception-V3 and VGG-16 models, while the Inception-V3 and VGG-16 models do not perform equally.

In the high-order potential function part, the Vision transformer model is selected as the feature extraction part of the high-order potential function of the whole model in this experiment. The difference between unary potential and binary potential is that the data used by the high-order potential is image-level data, that is, it contains the spatial attributes of the entire image. This applies to a wider range of spatial context information.

4.3 Experimental Results and Analysis of High Order Conditional Random Fields

The image-level classification results of three cervical cancer histopathological image datasets (AQP, HIF, VEGF) are depicted in Fig. 5 as a confusion matrix. This matrix represents the overall performance of the proposed high-order conditional random field model on the test set. By examining the confusion matrix, a clear assessment of the classification performance of the model can be obtained.

As can be seen from the confusion matrix in Fig. 5, the accuracy of the constructed high-order conditional random field model was 79.2%, 84.3% and 74.5%, respectively, on the verification sets of AQP, HIF and VEGF staining methods. This accuracy is significantly improved for the classification performance of the original unary potential, binary potential and high-order potential, and the classification accuracy of the test set is 1% to 9% lower than that of the verification intensification.

4.4 Classification Error Analysis

In the experimental results, some high-order conditional random field models misclassified the six experimental data sets. Through observation and analysis of experimental data, it can be speculated that the reasons for image classification errors are as follows:

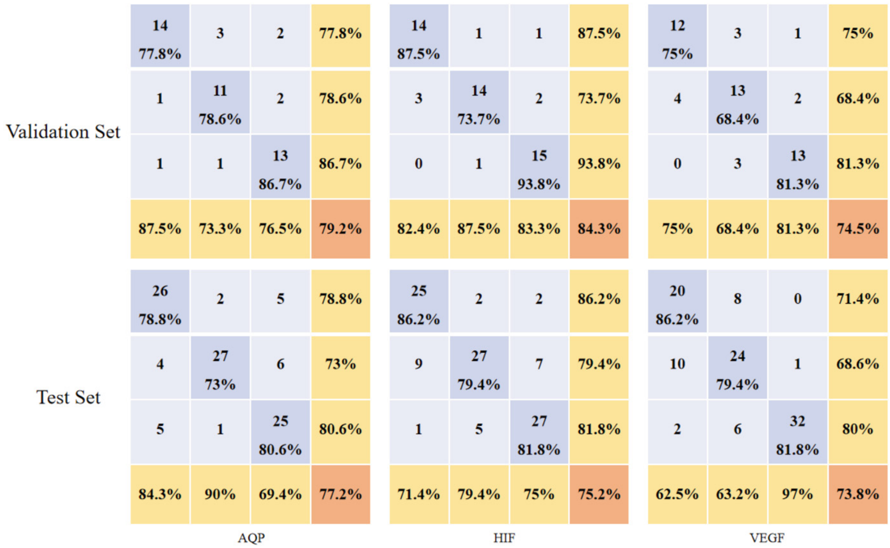

Fig. 5. The confusion matrix of the entire framework of the test set.

(1) First of all, because the internal information of the histopathological images of cervical cancer is very complex, and the image features and properties of each stage of differentiation are not always significantly different, it is difficult to extract the distinctive features. For example, the images of cervical cancer cells at the medium-differentiated stage are between the differentiated stages of cervical cancer cell images at the high-differentiated stage and low-differentiated stage, so it is often difficult to distinguish their features.

(2) Secondly, this paper uses a classification framework of weakly supervised learning. This method is convenient for the pathologist to work more efficiently, but the labeling information in the image data is crude and even inaccurate, because an image may contain information about different stages of differentiation. Therefore, errors in classification will be caused during model training. Figure 6 below is an example of an image test set misclassified on this model.

4.5 Classification Time

The training time of classification models for the three cervical cancer histopathological data sets is shown in Table 2 below.

As can be seen from Table 2, for AQP-stained data set, the average time to test an image is 2.31 s. For the HIF-stained dataset, the average time to test an image was 2.49 s; For the VEGF stained data set, an image was tested for 3.51 s. According to the test time, the proposed high order conditional random field model can be applied in clinical practice.

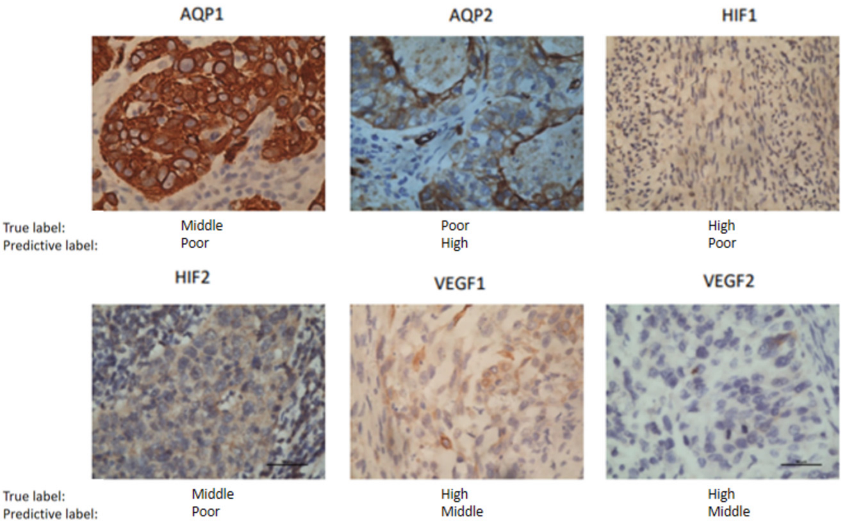

Fig. 6. A typical example of a misclassification result.

Table 2. Classification model training time on three cervical cancer histopathology datasets.

Dataset	Total time	Number of images	Mean time
AQP	228.73	99	2.31 s
HIF	261.32	105	2.49 s
VEGF	361.22	103	3.51 s
average	—	—	2.77 s

4.6 Visualization Analysis of High Order Conditional Random Field Model

To better demonstrate the effectiveness of the high-order conditional random field model in classifying high, medium, and low differentiation of cervical cancer histopathological images, this section compares its performance with classical deep learning models such as VGG-16, Inception-V3, ResNet-50, and the novel Vision Transformer model on the AQP, HIF, and VEGF datasets. The classification performance will be presented in a bar chart, with accuracy as the index. The index used is accuracy, and its comparison is shown in Fig. 7.

It can be seen that the high-order conditional random field model proposed in this experiment has better classification results for the high, medium and low differentiation classification of AQP, HIF and VEGF stained cervical cancer histopathological images, which has been improved in classification performance, proving the effectiveness of this model.

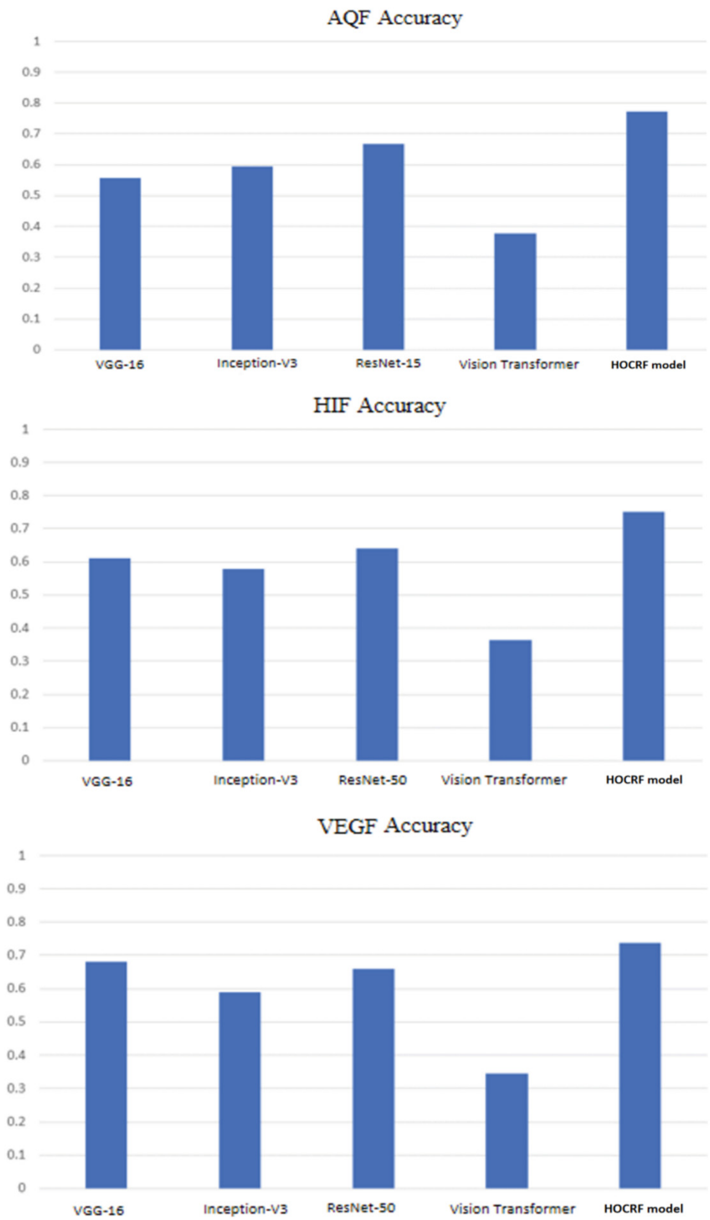

Fig. 7. Comparison and analysis of classification results of each model in histogram.

5 Conclusion and Future Work

This paper presents a novel approach for classifying high, medium, and low differentiation of cervical cancer histopathological images using a high-order conditional random field model with weak supervision. The model incorporates well-known deep learning models such as VGG-16, Inception-V3, and ResNet-50, along with a novel deep learning model called Vision Transformer. The proposed framework establishes the unary potential, binary potential, and high-order potential within the high-order conditional random field to capture the spatial relationship between image organization and cell position. Following the training and testing phases on the dataset, the model attains an overall accuracy of 75.4%. Additionally, the experiment incorporates visual analysis of the deep learning models. Comparative experiments demonstrate that the proposed high-order conditional random field model outperforms mainstream deep learning models, indicating the effectiveness of this approach. Future studies may explore combining traditional machine learning features with deep learning classifiers, as well as experimenting with different combinations of features and classifiers to further improve and optimize the model.

Acknowledgements. This work is supported by "National Natural Science Foundation of China" (No. 82220108007). We thank B.A. Yingying Hou from Foreign Studies College of Northeastern University, China, for her professional English proofreading in this paper. Xintong Li has the same important contribution to this paper as the first author. Chen Li has the same important contribution to this paper as the corresponding author.

References

1. Yang, X., Huang, L., Li, C., et al.: The effectiveness of mindfulness-based stress reduction intervention on alleviating anxiety and depression in postoperative patients with cervical cancer: A protocol for systematic review and meta-analysis. Medicine **101**(5) (2022)
2. Shin, H.R., Bray, F., et al.: Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int. J. Cancer **127**(12), 2893–2917 (2010)
3. Shrestha, A.D., Neupane, D., et al.: Cervical cancer prevalence, incidence and mortality in low and middle income countries: a systematic review. Asian Pacific J. Cancer Prevent. APJCP **19**(2), 319 (2018)
4. Buda, A., Borghese, M., et al.: Neoadjuvant chemotherapy prior fertility-sparing surgery in women with FIGO 2018 Stage IB2 cervical cancer: a systematic review. Cancers **14**(3), 797 (2022)
5. Bedell, S.L., Goldstein, L.S., Goldstein, A.R., et al.: Cervical cancer screening: past, present, and future. Sexual Med. Rev. **8**(1), 28–37 (2020)
6. William, W., Ware, A., et al.: A review of image analysis and machine learning techniques for automated cervical cancer screening from pap-smear images. Comput. Meth. Programs Biomed. **164**, 15–22 (2018)
7. Gautam, S., Bhavsar, A., Sao, A.K., et al.: CNN based segmentation of nuclei in PAP-smear images with selective pre-processing//Medical Imaging 2018: Digital Pathology. Int. Soc. Opt. Photon. **10581**, 105810X (2018)

8. Harangi, B., Toth, J., Bogacsovics, G., et al.: Cell detection on digitized pap smear images using ensemble of conventional image processing and deep learning techniques. In: 2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA), pp. 38–42. IEEE (2019)
9. Zhao, J., Dai, L., Zhang, M., et al.: PGU-net+: progressive growing of U-net+ for automated cervical nuclei segmentation. In: Li, Q., Leahy, R., Dong, B., Li, X. (eds.) Multiscale Multimodal Medical Imaging: First International Workshop, MMMI 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13, 2019, Proceedings, pp. 51–58. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-030-37969-8_7
10. Zhou, Y., Chen, H., Jiaqi, X., et al.: Irnet: Instance relation network for overlapping cervical cell segmentation. In: Shen, D., et al. (eds.) Medical Image Computing and Computer Assisted Intervention – MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part I, pp. 640–648. Springer International Publishing, Cham (2019). https://doi.org/10.1007/978-3-030-32239-7_71
11. Zhang, L., Lu, L., Nogues, I., et al.: DeepPap: deep convolutional networks for cervical cell classification. IEEE J. Biomed. Health Inform. **21**(6), 1633–1643 (2017)
12. Nanni, L., Ghidoni, S., Brahnam, S.: Handcrafted vs. non-handcrafted features for computer vision classification. Pattern Recogn. **71**, 158–172 (2017)

Skin Cancer Image Identification Using Deep Convolutional Neural Networks

Wenbo Zhang[1,2], Xiaolin Sun[1,2], Shiyi Ma[1,2], Jinzhu Yang[1,2], Marcin Grzegorzek[3,4], Yuexi Wang[5(✉)], and Chen Li[1,2]

[1] Microscopic Image and Medical Image Analysis Group, College of Medicine and Biological Information Engineering, Northeastern University, Shenyang, China
[2] Key Laboratory of Intelligent Computing in Medical Image, Ministry of Education, Northeastern University, Shenyang, China
[3] Institute of Medical Informatics, University of Luebeck, Luebeck, Germany
[4] Department of Knowledge Engineering, University of Economics in Katowice, Katowice, Poland
[5] Department of Rehabilitation, Shengjing Hospital of China Medical University, Shenyang, China
seabeta@163.com

Abstract. Skin cancer is a type of cancer and has emerged as one of the most prevalent types over the past decade. Early diagnosis and treatment are a key factor in the treatment of skin cancer. The traditional clinical diagnosis method of skin cancer is dermoscopy. On account of the complexity of the skin image, occasional instances of missed diagnosis and misdiagnosis arise, leading to delayed optimal treatment for patients. Two different networks, Inception-V3 and ResNet-50, are put forward in this paper. The dataset is separated into two categories, and the network performance is evaluated using evaluation indicators. Findings indicate that the ResNet50 network attains a higher accuracy rate of 88.83% compared to the Inception-V3 network's 83.17% accuracy. Additionally, the other three indicators of ResNet-50 are also surpass those of Inception-V3.

Keywords: Skin cancer · Image identification · Deep convolutional neural networks · Inception-V3 · ResNet-50

1 Introduction

Diseases affecting the skin and its appendages are known as dermatoses, which cause different degrees of changes in the morphology of the skin. According to clinical symptoms and morphological characteristics of the affected area, skin cancer can be divided into non-melanoma cell carcinoma and melanoma cell carcinoma. Non-melanoma cell carcinoma is more common, and melanocytic carcinoma has a higher fatality rate. According to the 2020 statistical report data of the American Cancer Society (AICR), melanoma skin cancer patients comprise only 21% of the total skin cancer cases in the world, but the mortality rate among those who succumb to malignant melanoma is alarmingly high at 17.6%, accounting for more than 47% of all skin cancer deaths [1].

P. You et al. (Eds.): ICIVIS 2023, LNEE 1163, pp. 17–27, 2024.
https://doi.org/10.1007/978-981-97-0855-0_2