

Research Article

Representation of Differential Learning Method for Mitosis Detection

Haider Ali ¹, Hansheng Li ¹, Ephrem Afele Retta ¹, Imran Ul Haq ¹,
Zhenzhen Guo ², Xin Han ¹, Lei Cui ¹, Lin Yang¹ and Jun Feng ¹

¹School of Information Science and Technology, Northwest University, Xian 710127, China

²Department of Thyroid Breast Surgery, Xi'an Daxing Hospital, Xi'an, Shaanxi, China

Correspondence should be addressed to Lei Cui; leicui@nwu.edu.cn

Received 12 November 2020; Revised 25 January 2021; Accepted 6 February 2021; Published 17 February 2021

Academic Editor: Jinshan Tang

Copyright © 2021 Haider Ali et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The breast cancer microscopy images acquire information about the patient's ailment, and the automated mitotic cell detection outcomes have generally been utilized to ease the massive amount of pathologist's work and help the pathologists make clinical decisions quickly. Several previous methods were introduced to solve automated mitotic cell count problems. However, they failed to differentiate between mitotic and nonmitotic cells and come up with an imbalance problem, which affects the performance. This paper proposes a Representation Differential Learning Method (RDLM) for mitosis detection through deep learning to detect the accurate mitotic cell area on pathological images. Our proposed method has been divided into two parts: Global bank Feature Pyramid Network (GLB-FPN) and focal loss (FL). The GLB feature fusion method with FPN essentially makes the encoder-decoder pay attention, to further extract the region of interest (ROIs) for mitotic cells. On this basis, we extend the GLB-FPN with a focal loss to mitigate the data imbalance problem during the training stage. Extensive experiments have shown that RDLM significantly outperforms on visualization view and achieves the best performance in quantitative matrices than other proposed approaches on the MITOS-ATYPIA-14 contest dataset. Our framework reaches a 0.692 F1-score. Additionally, RDLM achieves 5% improvements than GLB with FPN in F1-score on the mitosis detection task.

1. Introduction

According to the World Health Organization (WHO), over 2 million new cases were reported in 2018; furthermore, worldwide incidence and mortality ratios were 11.6% and 6.6% [1]. Breast cancer is the second most common cancer globally, usually in women in both developed and underdeveloped countries. It has been considered that breast cancer occurrences approximately exceed more than 50% from 2011 to 2030 [2]. As stated to the histologic tumour grade also named the Nottingham Grading System (NGS) from various expert global institutions, for instance, WHO, American Joint Committee on Cancer (AJCC), the Royal College of Pathologists (UK RCPATH), College of American Pathologists (CAP), and the European Union (EU) [3], NGS has consisted of three morphological features including nuclear pleomorphism, tubule formation, and mitotic index,

which are the most challenging part of breast cancer analysis. Each characteristic is given a score from 1 to 3. The ultimate NGS lowest possible score is 3 and highest possible score is 9, and NGS is divided into three grades: Grade 1 has been assigned for a total score of 3–5, also named well-differentiated; Grade 2 has been given for a total score of 6–7, also called moderately differentiated; and Grade 3 has been assigned for a total score of 8–9, also named poorly differentiated [4].

Automatic mitosis detection is a challenging task in microscopy images of breast cancer, especially in clinical practice. Detecting mitosis from stained high-power fields (HPFs), deal with difficulties such as whether the mitosis is in one of the four main phases; prophase, metaphase, anaphase, and telophase during mitosis development, which increases the complexity of the detection task as shown in Figure 1. Moreover, every phase has a very disparate shape

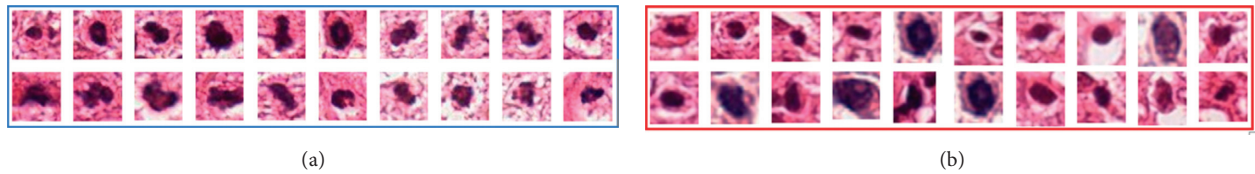


FIGURE 1: (a) Blue rectangle represents the actual mitotic cells, and (b) red rectangle represents the nonmitotic cells, which have a similar morphological appearance.

and texture configurations in four phases. For instance, a nucleus in telophase is split into two distinct regions, even though they are still one connected mitotic cell. Nonmitotic cells such as apoptotic, dense nuclei, and lymphocytes similar to mitotic cells on the slides have a similar morphological appearance. The dealing slide obtainment can present artefacts and undesirable objects, distinguishing cells difficultly and leading to many false positives in the detection process [5]. In the last few years, various efficient automatic mitosis systems have developed to detect mitotic cells on pathological images. Automatic mitosis detection has been a very active research area because of the recent advancement in digital medication, especially helping pathologists to make clinical decisions. Some machine learning handcrafted features approaches [6–12] have been used to extract the features. However, these traditional approaches always suffer from the large shape variations of mitosis and also need a very much high exertion time to approve and cannot signify characteristics or features adequately of mitosis regions, resulting in low performance on the different contest dataset such as ICPR 2012 [13] contest, AMIDIA 2013 [14] contest, and ICPR 2014 [15] contest datasets.

Computer vision is one of the areas advancing rapidly [16]. Deep learning (DL) computer vision has helped to ease the massive amount of pathologists' work, and the pathologists are making fast clinical decisions. DL also makes mitosis detection work much easier than before such as in mitotic cell classification task [17–20], mitotic cells detection task [21, 22], and mitotic cells segmentation task [1, 23–25]. Deep learning techniques outperform preliminary research techniques, for instance, handcrafted features that capture specific mitosis features for automatic detection. Deep learning techniques are based on CNN and a deep classifier like Deep Cascade Neural Network (CasNN). CNN is first used to extract candidate regions, and a deep classifier is used to differentiate cancer and noncancer cells [17]. Deep Mitosis (DM) technique, which consists of segmentation task, detection task, and classification task [21], always fails to differentiate between mitosis and nonmitosis cell shape. Recently, the object detection and instance segmentation framework Feature Pyramid Network (FPN) have been used to tackle different kinds of medical problems, such as detecting the porosity and cracks in concrete CT images [26], automatic segmentation of cervical nuclei [27], thyroid nodule detection from medical ultrasound images [28], detection of teeth and their components in X-ray images [29], lung nodule detection in chest X-rays images [30], and having attained outstanding results. The focal loss (FL) has been used to solve the data imbalance problem in the various

kinds of biomedical datasets and performed well for minority class, such as classification of red blood cells morphology [31], colon gland instance segmentation [32], and localization of cell organelles [33]. The mitoses are embedded in complex backgrounds and influenced by various factors containing, staining, lightning condition, tissue acquisition, and similar appearance of nonmitosis. In 2019, the global bank (GLB) feature fusion method was proposed to tackle the complex background problem on the 2015 Gland dataset. This feature fusion method achieved outstanding results on both segmentation and cancer embolus detection tasks compared to other proposed approaches [34].

Inspired by the above object detection framework and the GLB feature fusion method, we considered the mitosis detection task as an object detection problem. We proposed a new method for mitosis detection through deep learning to detect the mitotic cell area on a pathological image, called RDLM, which is fast and accurate to detect mitotic cells for both slide scanner Aperio-type and Hamamatsu-type images as shown in Figure 2. Our proposed framework is divided into two parts: GLB and FL. The GLB feature fusion method with FPN essentially calibrates the encoder-decoder and further makes encoder-decoder pay attention to extract the region of interest (ROIs) for mitotic cells. GLB has three phases. In the first phase, the encoder connects the layer at multiple-scales by including features. The second phase reconnects and removes the background noise to acquire a suitable feature map that further pays attention to the region of interest for mitotic cells. In the final phase, the feature map is generally transferred to the decoder by convolutions. Besides, GLB outperforms on the MITOS-ATYPIA 2014 contest dataset comprised of HPF large shaped variations of mitosis and the similar appearance of nonmitosis that creates severe problems during the training phase. However, the training process cannot learn full information on positive samples because most of them are easy and simple disguisable negative samples, bringing the imbalance problem. We adopted the FL to alleviate this scenario and obtained good mitosis detection results on both slide scanners. To validate the effect of our proposed method RDLM in visualization view and quantitative matrices, we conducted experiments on the MITOS-ATYPIA 2014 contest dataset. Extensive experiments show that RDLM attained outstanding results compared to other state-of-the-art methods on both slide scanners.

This research's main contributions could be summarized as follows: first, we introduced the feasibility and superiority of a global bank and focal loss in mitosis detection tasks. Second, we proposed a new method called RDLM for mitosis detection through deep learning to detect the mitotic cell

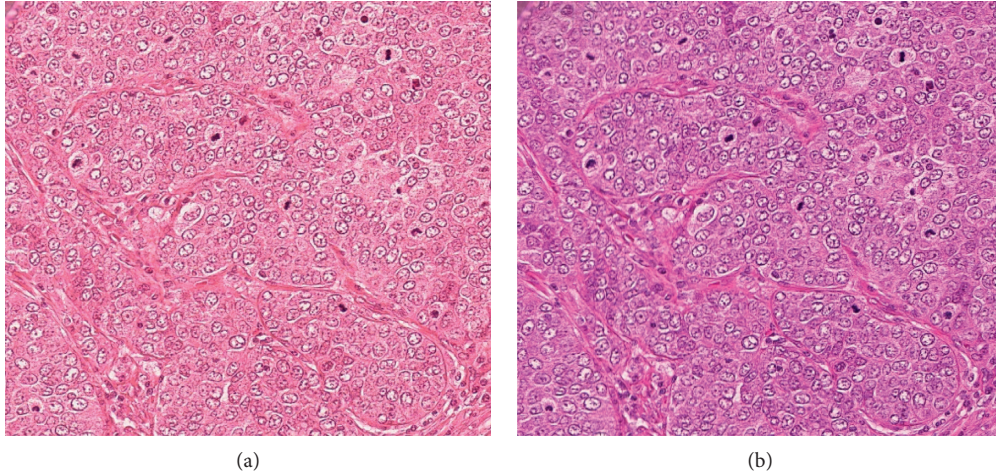


FIGURE 2: Example of ICPR2014 contest dataset two slide scanner HPF images: (a) taken from Aperio-type scanner and (b) taken from the Hamamatsu-type scanner.

area on pathological images. The GLB makes encoder-decoder to pay attention further to extract the ROIs for mitotic cells while the FL alleviates the imbalance problem resulting from the predicted detection of mitotic cells results. Third, we applied extensive experiments to validate the effectiveness of the proposed method. The results showed that our proposed method attained the best performance in mitosis detection tasks on both slide scanners in terms of visualization view and quantitative metrics compared to other state-of-the-art approaches.

2. Methodology Motivation

This section describes our proposed method for mitosis detection through deep learning to detect the mitotic cell area on a pathological image, as shown in Figure 3. The FPN framework builds feature pyramids inside the convolution neural network (CNN). Moreover, FPN delivers a top-down pathway to produce higher-resolution layers from a semantic rich layer. The GLB feature fusion method part essentially calibrates the encoder-decoder and makes the encoder-decoder pay attention further to extract the ROIs for the mitotic cells detection task. We adopted the FL to alleviate the imbalance problem during the training stage, and each component's details as described in the following sections.

2.1. FPN. Feature Pyramid Network is a newer, clean, and simple framework for building feature pyramids inside the CNN. It delivers a top-down pathway to produce higher-resolution layers from a semantic rich layer. Simultaneously, the region proposal network (RPN) used a sliding subnetwork at every location over the multilevel feature maps to generate an object proposal. An anchor defines every object proposal, and each anchor scale has a corresponding level in the features pyramid [35]. Besides, CNN has been employed to extract each patch's comprehensive features from the input image. We used ResNet-50, which is the most efficient and powerful in image classification. The WSI sampled

patches are fixed size 1040×1040 pixels, so we have resized the patches to 600×600 pixels and then fed them into the FPN.

2.2. GLB

2.2.1. The Vault Layer. In short, the vault layer stores the well-chosen features compared to other layers in the equal phase. In the encoding phase, various consecutive layers construct the result maps of the equal dimension that are reckoned in a similar network phase. In a phase of encoder, we represent the last layer by $\{EN_i, i \in [1, s]\}$, where s denote the encoder number of phases. EN_i extracts better-quality features than different layers extracted features in a similar phase. By accumulating all EN_i , we can link the layers of phases in the encoder. Though, every EN_i has a unique dimension and depth, before accumulating them, the dimension of EN_i to upsample them and later utilize $3 \times 3 \times k$ convolutions, where k denote hyperparameter. Finally, we include an entire EN_i elementwise and $3 \times 3 \times m$ convolutions utilized to specifically store the accumulated features in the vault layer, where m is a hyperparameter. The vault layer equation can be described as

$$VL = f_{3 \times 3}^A \left[f_{3 \times 3}^{EN1}(EN1) + \sum_{i=2}^s f_{3 \times 3}^{ENi}(\text{up}(EN_i)) \right], \quad (1)$$

where $f_{3 \times 3}^A$ represents $3 \times 3 \times m$ convolutions. $f_{3 \times 3}^{ENi}$ for $i \in [1, s]$ represents $3 \times 3 \times k$ convolutions corresponding to EN_i , where s denotes the encoder number of phases. up describes upsampling of EN_i to the dimension of EN_1 .

2.2.2. Calibration Layers and Gain Layers. In general, for detection tasks, a calibration layer is denoted by $\{\text{CAL}_i, i \in [1, s]\}$, where s denotes the decoder number of phases during decoding phases. Every calibration layer has an altered resolution and depth. The vault layer extracted features maps refer to gain layers and represent them

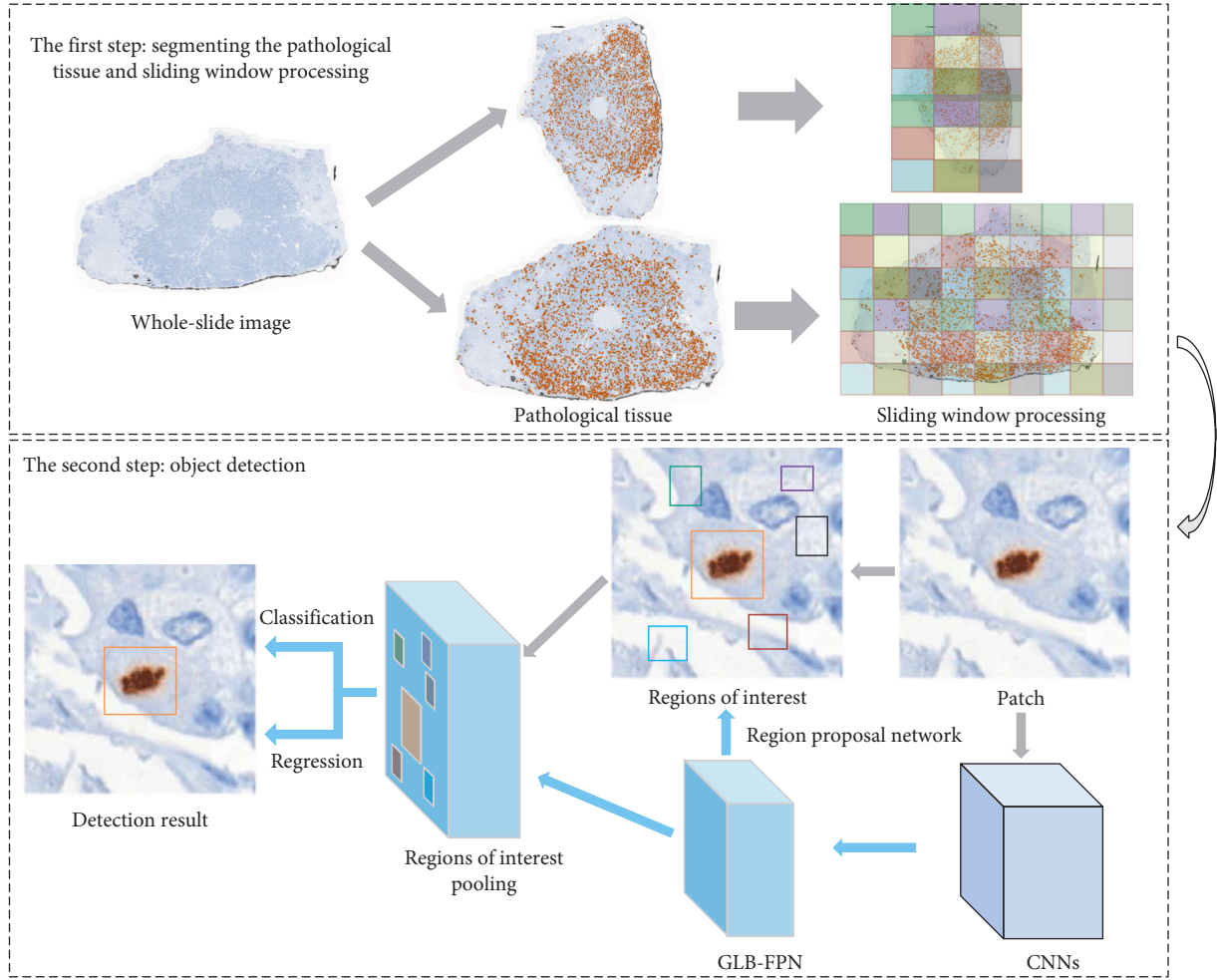


FIGURE 3: A simple overview of our RDLM method for mitotic cell detection on pathological images.

through $\{GAIN_i, i \in [1, s]\}$, where s is the number of CAL_i . At that point, CAL_i can be enhanced and calibrated with comparing $GAIN_i$. We equalize the depth by using 3×3 convolutions with a stride of 2 and downsampling to the dimension of CAL_i as illustrated in equation (2). In one case, the downsampling operation is avoidable when the calibration layer's dimension is equivalent to that of the vault layer and is illustrated in the following equation:

$$GAIN_i = \begin{cases} f_{GAIN_i}(VL), \\ \text{down}(f_{GAIN_i}(VL)), \end{cases} \quad (2)$$

where f_{GAIN_i} denotes 3×3 convolutions with a stride of 2 and downsampling operations utilized to equalize the dimension of features with that of CAL_i . At that point, CAL_i can be enhanced and calibrated with comparing $GAIN_i$. So, we include calibrated and gain layers as elementwise, as illustrated in the following equation:

$$CAL_i = \{CAL_i \oplus GAIN_i\}. \quad (3)$$

The design of the proposed GLB with FPN is shown in Figure 4, where we first include elementwise calibration layers (CAL_i) and gain layers ($GAIN_i$) as illustrated in

equation (3). Using a gain layer to enhance a predicting layer permits the full attention of local features, powerful rich semantic features, and the gain layer global features, resulting in improved predictions.

2.3. An Objective of Focal Loss. In 2017, the focal loss was proposed by Kaiming He's team [36]. The author uses a piecewise function to represent the cross-entropy of binary classification problem between foreground and background classes, and a piecewise function of p_t to represent y_i with the value of 1 in binary classification. The value here is the category of 1 in one-hot [0 1] or [1 0]:

$$CE(p, y) = \begin{cases} -\log(p), & \text{if } y = 1, \\ -\log(1 - p), & \text{otherwise,} \end{cases} \quad (4)$$

$$P_t = \begin{cases} p, & \text{if } y = 1, \\ 1 - p, & \text{otherwise,} \end{cases}$$

where p_t denotes the model prediction probability and y indicates the ground truth label. Hence, $CE(p, y) = CE(p_t) = -\log(p_t)$, CE with a balancing factor, where α denotes separate categories:

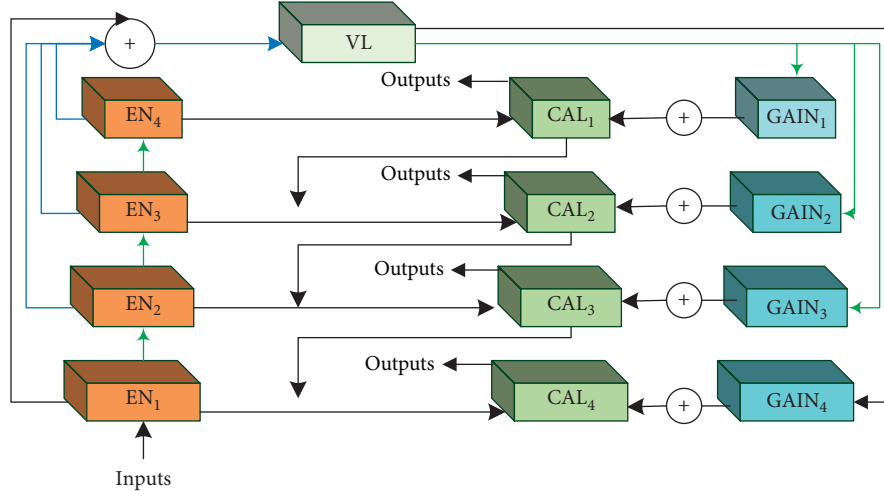


FIGURE 4: Sketch of the GLB feature fusion method consists of three layers: VL layers, gain layers, and calibration layers [34]. The green line denotes the Conv + down, the blue line represents the Conv + up, and the grey line denotes the Conv.

$$CE(p_t) = -\log(p_t) = -\alpha_t \log(p_t). \quad (5)$$

So, the γ factor is introduced to focus more on the hard negatives samples during training. Thus, a focal loss is defined as

$$FL = -\alpha_t (1 - p_t)^\gamma \log(p_t). \quad (6)$$

In our case, we have integrated the focal loss to solve the data imbalance problem in the mitosis detection task in our proposed method. The multilabels focal loss formula is defined as

$$FL = -\alpha_t * (z - p_t)^\gamma * \log(p_t) - (1 - \alpha_t) * p_t^\gamma * \log(1 - p_t), \quad (7)$$

where alpha is equal to 0.5, gamma is equal to 0.5, p_t is equal to sigmoid(x), and z is equal to the target in equation (7). The prediction p_t represents the predicted logits for each class and the target z represents the one-hot encoded classification targets. α and γ are hyperparameters. For the positive prediction, we only need to consider the front part loss, and the back part is 0. z has been considered as greater than zero, less than or equal to zero, and z is equal to 1, so the positive coefficient is equal to $z - p_t$. For the negative prediction, we only need to consider back part loss, and the front part is 0; z has been considered as greater than zero, less than or equal to zero, and z is equal to 1, so the negative coefficient is equal to 0.

3. Experiments

We have conducted experiments based on the RDLM method to examine the impact of accurate mitotic cell detection. Besides, we used the FPN baseline, FPN-FL, and GLB-FPN on both slide scanners. Finally, we obtained

improved results on the RDLM proposed method, verifying this research work's success.

3.1. Implementations Details. Our approach is implemented with python using Tensorflow libraries and tested on a machine with NVIDIA GeForce GTX 1080 Ti GPU. It takes 4.5 hours for training images and 2.3 minutes for testing images. ImageNet pretrained models ResNet-50 [18] has been used as our base network for FPN. The input image is resized such that its shorter side has 600 pixels. We adopt the SGD for optimization, set momentum as 0.9, and set weight decay as 0.0001. The learning rate has been set at 0.001 for 20k iteration steps.

3.2. Dataset. We have conducted experiments on the MITOS-ATYPIA-14 contest dataset. The contest data samples were scanned by two slide scanners Aperio Scanscope XT and Hamamatsu Nanozoomer 2.0-HT, and the whole-slide histological images (WSIs) were stained with standard hematoxylin and eosin (H&E) dyes. The training set of Aperio Scanscope XT scanner data at X40 magnification provided 1,136 labelled 1539 × 1376 pixels HPFs frames, 749 labelled mitotic cells, and 496 are unlabeled. The centroids pixels of mitoses have been manually annotated via two senior pathologists. In a situation of contradiction between the pathologists, the 3rd one will provide the last say. We have divided the training dataset of Aperio-type into two parts; 300 images used for training, and 92 used for testing in our experiments.

3.3. Evaluation Metrics. We adopt the same contest evaluation criteria of the MITOSIS ICPR challenges; correct detection would be considered in this case if the distance to a ground truth candidate mitosis is less than 8 μm . F1-score,

TABLE 1: Mitosis detection performance with the proposed method on the Aperio-type images.

Method	α	γ	k	m	Precision	Recall	F1-score
FPN baseline	—	—	—	—	0.578	0.599	0.589
FPN-FL	0.5	0	—	—	0.586	0.613	0.595
FPN-FL	0.5	0.5	—	—	0.614	0.645	0.629
FPN-FL	0.5	1	—	—	0.563	0.584	0.573
FPN-FL	0.5	2	—	—	0.496	0.647	0.561
GLB-FPN	—	—	256	256	0.614	0.675	0.643
GLB-FPN	—	—	256	512	0.647	0.685	0.662
RDLM	0.5	0	256	256	0.623	0.654	0.638
RDLM	0.5	0.5	256	256	0.685	0.70	0.692
RDLM	0.5	1	256	256	0.591	0.640	0.614
RDLM	0.5	2	256	256	0.546	0.674	0.602

TABLE 2: Comparison with other state-of-the-art approaches on the Aperio-type images.

Method	Precision	Recall	F1-score
STRASBOURG	—	—	0.024
YILDIZ	—	—	0.167
MINES-CURIE-INSERM	—	—	0.235
CUHK	0.448	0.300	0.356
DeepMitosis [21]	0.431	0.443	0.437
CasNN [17]	0.411	0.478	0.442
MaskMitosis [37]	0.500	0.453	0.475
LRCNN + in group [38]	0.654	0.663	0.659
Efficient mitosis detection [39]	0.534	0.661	0.585
SegMitosis-r15R30 [5]	0.594	0.512	0.550
SegMitosis-random [5]	0.637	0.502	0.562
RDLM	0.685	0.70	0.692

precision, and recall are used to compute the performance evaluation of our method:

$$F1 - score = 2 * \frac{(\text{precision} * \text{recall})}{(\text{precision} + \text{recall})},$$

$$\text{precision} = \frac{TP}{(TP + FP)}, \quad (8)$$

$$\text{recall} = \frac{TP}{(TP + FN)}.$$

The TP, FP, and FN are the number of true positive detections, the number of false-positive detections, and the number of false-negative, which means they are undetected.

3.4. Results. Table 1 shows the proposed method RDLM performance for mitosis detection through deep learning quantitative results on the Aperio-type slide scanner. The first part of the table lists the FPN baseline, FPN-FL, GLB-FPN, and RDLM results. The next four parts of the table represent the hyperparameters; α , γ , k , and m , where k represents the convolutional filters from EN_i , m represents the vault layer depth, α is an imbalance factor, and γ focuses on the hard negatives examples in our experiments on the training set. Lastly, the last three parts of the table list the evaluation metrics results. As we can observe from Table 1

when α is 0.5, γ is 0.5, k is 256, and m is 256, our proposed method achieves the best performance F1-score 0.692 compared to other methods. The number of parameters in the RDLM is more than other methods of GLB-FPN and FPN-FL. Instead of the fact that adding more parameters makes training more complicated in the network architecture, the entire performance of the RDLM is better than that of the FPN baseline, FPN-FL, and GLB-FPN.

3.5. Discussion. The mitosis cell detection for the MITOS-ATYPIA-14 contest dataset is one of the most challenging compared to other datasets like MITOSIS 2012, AMEDIA 2013, and TUPAC 2016 due to more complicated background tissue appearance and a variation of shape. Besides, the samples of this dataset have been weakly annotated. However, several methods have been proposed, and they cannot obtain the best F1-score from the challenges discussed above on this dataset. According to the experimental evaluation, as shown in Table 1, it can be noted that our proposed method achieved the best performance result with F1-score of 0.692 compared with current deep learning mainstream approaches in Table 2. The first four methods results have been taken from the MITOS-ATYPIA-14 contest, and ‘—’ denotes the not released results. Later, several other CNN based methods have published. The DeepMitosis [21] and MaskMitosis [37] yield a good performance on the

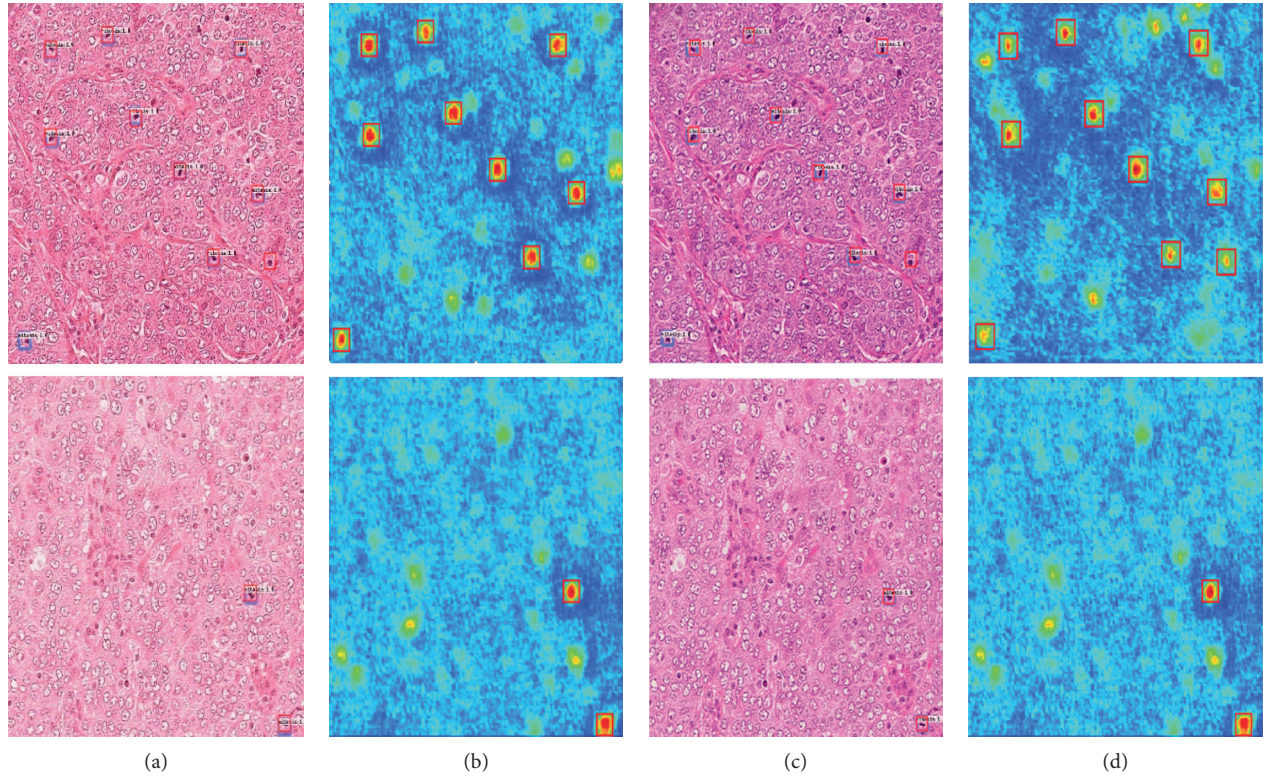


FIGURE 5: Detection results of examples on the testing set from both slide scanners.

MITOSIS 2012 dataset but inferior performance on the MITOS-ATYPIA-14 contest dataset. These two segmentation modules were not reliable enough during the detection phase and created inferior performance on the weakly annotated dataset. The CasNN [17] approach requires two different networks: one is used to retrieve the mitosis candidates, and the other is used to classify the candidates, leading to less accurate detection. The LRCNN [38] and SegMitosis [5] approaches reported F1-score of 0.659 and 0.562. In contrast, the proposed method attained better performance than existing methods on the testing set in terms of quantitative matrices F1-score of 0.692.

4. Visualization and Discussion

In this section, we will discuss the qualitative results of mitosis detection. Furthermore, we analyze the effect of FL in visual view to make the predicted results more reasonable.

4.1. Visualization of Mitotic Cells. The visualizations of mitotic cells detection have been done in RDLM, where GLB is converting them into a heatmap. The vault layer focuses more on the HPF tissue regions with Aperio-type and Hamamatsu-type scanners images, and the gain layer focuses more on the inside area of the Aperio-type and Hamamatsu-type scanners images. Moreover, FL makes the result more sensible from predicted results, as shown in Figure 5. The equation of heatmap is defined as

$$HM = A \times \alpha + (1 - \alpha) \times \text{Image}. \quad (9)$$

HM represents the heatmap. A means the activation function. α represents weight coefficient, and the image represents the feed input images. Figure 5 describes the results of the mitosis detection examples on the ICPR 2014 dataset where samples were scanned with both slide scanners, left side (a), (b) denote Aperio-type, and right side (c), (d) denote Hamamatsu-type. (a) and (c) are the predictions of the RDLM ($k=256$, $m=256$, $\alpha=0.5$ and $\gamma=0.5$) results on both slides scanner, where red labels denote the ground truth and blue is the model prediction. (b) and (d) are the visualization results on both slide scanners where manual labels are red.

5. Discussion

This section explains the parameter selection of focal loss from two aspects: quantitative evaluation and visual perception. First, we can notice from Table 1 that the quantitative evaluation of FL decreases the probability of false-positive predictions and false-negative predictions to make the predicted results more reasonable when α is 0.5 and γ is 0.5. We set α to 0.5 with different γ of 0, 0.5, 1, and 2, respectively. However, the RDLM model performs well compared to other methods in terms of F1-score metrics and achieves the highest score on Aperio-type test images. Second, to illustrate the visual effects of FL, we selected three examples from the Aperio-type slide scanner's testing set. We can observe from

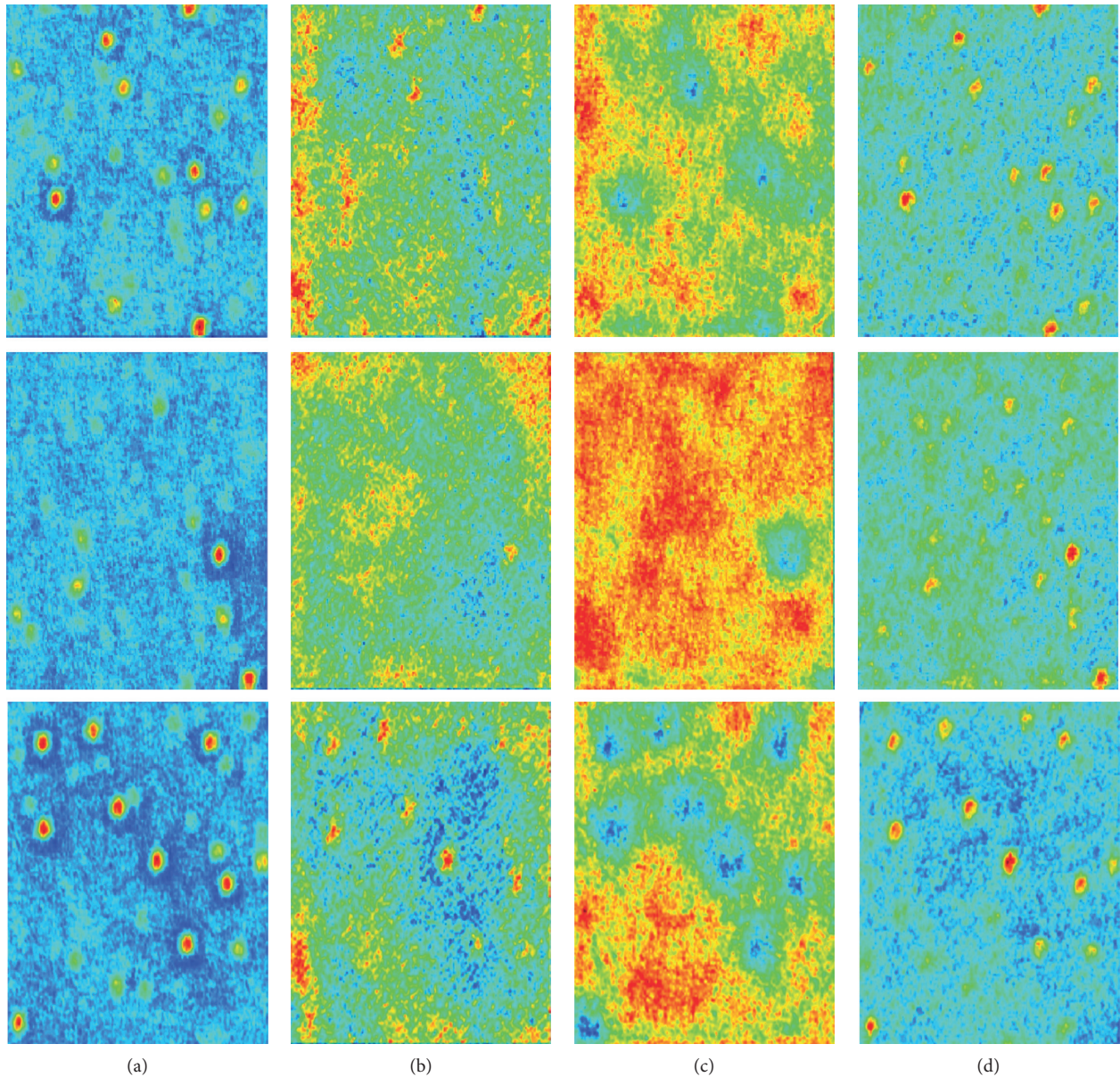


FIGURE 6: visual effects on imbalance factor $\alpha=0.5$ with various focusing factors γ . (a) $\alpha=0.5$, $\gamma=0.5$. (b) $\alpha=0.5$, $\gamma=0$. (c) $\alpha=0.5$, $\gamma=1$. (d) $\alpha=0.5$, $\gamma=2$.

Figure 6 that the RDLM obtained good results when α is 0.5 and γ is 0.5. In contrast, the RDLM method poorly performs over all the examples when γ is 0, 1, 2, and the detection results always have some noise and misclassification regions of mitotic cells. Finally, we can conclude that the proposed method performs well and attained high performance in terms of both quantitative evaluation and visual perception when α is 0.5 and γ is 0.5.

6. Conclusion and Future Direction

In this paper, we proposed a new method for mitosis detection tasks through deep learning called RDLM. Our proposed framework has been divided into two parts: GLB-FPN and FL. The GLB feature fusion method with

FPN essentially calibrates the encoder-decoder and makes the encoder-decoder pay attention further to extract the ROIs for mitotic cells. FL is adopted to efficiently alleviate the data imbalance problem in mitotic cell detection tasks. Besides, extensive experiments were carried out on the MITOS-ATYPIA-14 contest dataset to verify the effectiveness of the proposed method. Results showed that the proposed approach is superior to most current mainstream approaches in visualization view and quantitative matrices. Compared with state-of-the-art techniques, our framework has achieved a higher score on the testing set in terms of quantitative metrics F1-score of 0.692. We intend to design a new pipeline for both detection and segmentation tasks of mitotic cells in the future study.

Data Availability

In our experiments, we have used the publicly available dataset MITOS-ATYPIA-14 contest (<https://mitos-atypia-14.grand-challenge.org/Download/>). Moreover, we have also cited this dataset in our references.

Conflicts of Interest

The authors would like to confirm that they have no conflicts of interest, financial or others.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (no. 62073260), the key problems in Medical Image Analysis with the Fusion of Vision Perception and Cognition.

References

- [1] M. Z. Alom, T. Aspiras, T. M. Taha, T. J. Bowen, and V. K. Asari, "MitosisNet: end-to-end mitotic cell detection by multi-task learning," *IEEE Access*, vol. 8, pp. 68695–68710, 2020.
- [2] M. C. A. Balkenhol, P. Bult, D. Tellez et al., "Deep learning and manual assessment show that the absolute mitotic count does not contain prognostic information in triple negative breast cancer," *Cellular Oncology*, vol. 42, no. 4, pp. 555–569, 2019.
- [3] A. T. Atanda, M. I. Imam, A. B. Umar et al., "Audit of nottingham system grades assigned to breast cancer cases in a Teaching Hospital," *Annals of Tropical Pathology*, vol. 8, no. 2, 2017.
- [4] E. A. Rakha, J. S. Reis-Filho, F. Baehner et al., "Breast cancer prognostic classification in the molecular era: the role of histological grade," *Breast Cancer Research*, vol. 12, no. 4, pp. 1–12, 2010.
- [5] C. Li, X. Wang, W. Liu, L. J. Latecki, B. Wang, and J. Huang, "Weakly supervised mitosis detection in breast histopathology images using concentric loss," *Medical Image Analysis*, vol. 53, pp. 165–178, 2019.
- [6] C. Sommer, L. Fiaschi, F. A. Hamprecht et al., "Learning-based mitotic cell detection in histopathological images," in *Proceedings of the 21st International Conference on Pattern Recognition*, pp. 2306–2309, Tsukuba Science City, Japan, 2012.
- [7] H. Irshad, "Automated mitosis detection in histopathology using morphological and multi-channel statistics features," *Journal of Pathology Informatics*, vol. 4, 2013.
- [8] A. M. Khan, H. El-Daly, and N. M. Rajpoot, "A gamma-Gaussian mixture model for detection of mitotic cells in breast cancer histopathology images," in *Proceedings of the 21st International Conference on Pattern Recognition*, pp. 149–152, Tsukuba Science, Japan, 2012.
- [9] C. D. Malon and E. Cosatto, "Classification of mitotic figures with convolutional neural networks and seeded blob features," *Journal of Pathology Informatics*, vol. 4, 2013.
- [10] C. H. Huang and H. K. Lee, "Automated mitosis detection based on exclusive independent component analysis," in *Proceedings of the 21st International Conference on Pattern Recognition*, pp. 1856–1859, Tsukuba Science, Japan, 2012.
- [11] F. B. Tek, "Mitosis detection using generic features and an ensemble of cascade adaboosts," *Journal of Pathology Informatics*, vol. 4, 2013.
- [12] A. Paul, A. Dey, D. P. Mukherjee, J. Sivaswamy, and V. Tourani, "Regenerative random forest with automatic feature selection to detect mitosis in histopathological breast cancer images," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 94–102, Munich, Germany, 2015.
- [13] L. Roux, D. Racoceanu, N. Loménié, M. Kulikova et al., "Mitosis detection in breast cancer histological images an ICPR 2012 contest," *Journal of Pathology Informatics*, vol. 4, p. 2013.
- [14] M. Veta, P. J. Van Diest, S. M. Willems et al., "Assessment of algorithms for mitosis detection in breast cancer histopathology images," *Medical Image Analysis*, vol. 20, no. 1, pp. 237–248, 2015.
- [15] MITOS-ATYPIA-14: Mitos-atypia-14-dataset, 2014, [https://mitos-atypia-14.grand-challenge.org/Dataset/\(2014\)](https://mitos-atypia-14.grand-challenge.org/Dataset/(2014)).
- [16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [17] H. Chen, Q. Dou, X. Wang et al., "Mitosis detection in breast cancer histology images via deep cascaded networks," in *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, Phoenix, AZ, USA, February 2016.
- [18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, 2016.
- [19] D. C. Cireşan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Mitosis detection in breast cancer histology images with deep neural networks," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 411–418, Berlin, Germany, 2013.
- [20] H. Wang, A. Cruz-Roa, A. Basavanthally et al., "Cascaded ensemble of convolutional neural networks and handcrafted features for mitosis detection," in *Proceedings of SPIE-The International Society for Optical Engineering*, Bellingham, WA, USA, 2013.
- [21] C. Li, X. Wang, W. Liu, and L. J. Latecki, "DeepMitosis: mitosis detection via deep detection, verification and segmentation networks," *Medical Image Analysis*, vol. 45, pp. 121–133, 2018.
- [22] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: towards real-time object detection with region proposal networks," in *Proceedings of the Advances in Neural Information Processing Systems*, pp. 91–99, Montreal, Canada, 2015.
- [23] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2961–2969, Venice, Italy, 2017.
- [24] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440, Boston, MA, USA, 2015.
- [25] K. S. Beevi, M. S. Nair, and G. R. Bindu, "Automatic mitosis detection in breast histopathology images using Convolutional Neural Network based deep transfer learning," *Bio-cybernetics and Biomedical Engineering*, vol. 39, no. 1, pp. 214–223, 2019.
- [26] L. Zhao, C. J. Li, X. D. Wu, D. T. Niu, Z. X. Duan, and F. N. Dang, "Improved damage characteristics identification method of concrete CT images based on region convolutional

- neural network,” *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 34, no. 6, p. 2054018, 2018.
- [27] Y. Liu, P. Zhang, Q. Song, A. Li, P. Zhang, and Z. Gui, “Automatic segmentation of cervical nuclei based on deep learning and a conditional random field,” *IEEE Access*, vol. 6, pp. 53709–53721, 2018.
- [28] S. Yao, J. Yan, M. Wu et al., “Texture synthesis based thyroid nodule detection from medical ultrasound images: interpreting and suppressing the adversarial effect of in-place manual annotation,” *Frontiers in Bioengineering and Biotechnology*, vol. 8, p. 2020.
- [29] G. Jader, J. Fontinel, M. Ruiz, K. Abdalla, M. Pithon, and L. Oliveria, “Deep instance segmentation of teeth in panoramic X-ray images,” in *Proceedings of the 2018 31st SIB-GRAPI Conference on Graphics*, pp. 400–407, Foz do Iguaçu, Brazil, 2019.
- [30] C. Liu, B. Wang, Q. Jiao, and M. Zhu, “Reducing false positives for lung nodule detection in chest X-rays using cascading CNN,” in *Proceedings of the 2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pp. 1204–1207, Xi’an, China, 2019.
- [31] K. Pasupa, S. Vatathanavaro, and S. Tungjitnob, “Convolutional neural networks based focal loss for class imbalance problem: a case study of canine red blood cells morphology classification,” *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–17, 2020.
- [32] L. Mei, X. Guo, X. Huang, Y. Weng, S. Liu, and C. Lei, “Dense contour-imbalance aware framework for colon gland instance segmentation,” *Biomedical Signal Processing and Control*, vol. 60, p. 101988, 2020.
- [33] N. Yudistira, M. Kavitha, T. Itabashi, A. H. Iwane, and T. Kurita, “Prediction of sequential organelles localization under imbalance using a balanced deep u-net,” *Scientific Reports*, vol. 10, no. 1, pp. 1–11, 2020.
- [34] H. Li, B. Kang, Y. Kang et al., “Global Bank: A guided pathway of encoding and decoding for pathological image analysis,” in *Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine*, pp. 415–422, San Diego, CA, USA, 2019.
- [35] T. Y. Lin, P. Dollár, R. Girshick et al., “Feature pyramid networks for object detection,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2117–2125, Honolulu, HI, USA, 2017.
- [36] T. Y. Lin, P. Goyal, R. Girshick et al., “Focal loss for dense object detection,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2980–2988, Venice, Italy, 2017.
- [37] M. Sebai, X. Wang, and T. Wang, “MaskMitosis: a deep learning framework for fully supervised, weakly supervised, and unsupervised mitosis detection in histopathology images,” *Medical & Biological Engineering & Computing*, vol. 58, pp. 1603–1623, 2020.
- [38] Y. Li, E. Mercan, S. Knezevitch, J. G. Elmore, and L. G. Shapiro, “Efficient and accurate mitosis detection—a lightweight RCNN approach,” in *Proceedings of the International Conference on Pattern Recognition Applications and Method*, pp. 69–77, Funchal, Portugal, 2018.
- [39] D. Cai, X. Sun, and N. Zhou, “Efficient mitosis detection in breast cancer histology images by RCNN,” in *Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pp. 919–922, Venice Italy, 2019.