

A Loopback Network for Explainable Microvascular Invasion Classification

Shengxuming Zhang¹, Tianqi Shi², Yang Jiang², Xiuming Zhang¹, Jie Lei³, Zunlei Feng^{1,4,5,*}, Mingli Song^{1,4,5}

¹Zhejiang University, ²Alibaba Group, ³Zhejiang University of Technology,

⁴Shanghai Institute for Advanced Study of Zhejiang University,

⁵Zhejiang Provincial Key Laboratory of Service Robot

{zsxm1998, 1508056, zunleifeng, brooksong}@zju.edu.cn,

{tianqi.tianqishi, yangjiang.yj}@alibaba-inc.com, jasonlei@zjut.edu.cn

Abstract

Microvascular invasion (MVI) is a critical factor for prognosis evaluation and cancer treatment. The current diagnosis of MVI relies on pathologists to manually find out cancerous cells from hundreds of blood vessels, which is time-consuming, tedious, and subjective. Recently, deep learning has achieved promising results in medical image analysis tasks. However, the unexplainability of black box models and the requirement of massive annotated samples limit the clinical application of deep learning based diagnostic methods.

In this paper, aiming to develop an accurate, objective, and explainable diagnosis tool for MVI, we propose a Loopback Network (LoopNet) for classifying MVI efficiently. With the image-level category annotations of the collected Pathologic Vessel Image Dataset (PVID), LoopNet is devised to be composed binary classification branch and cell locating branch. The latter is devised to locate the area of cancerous cells, regular non-cancerous cells, and background. For healthy samples, the pseudo masks of cells supervise the cell locating branch to distinguish the area of regular non-cancerous cells and background. For each MVI sample, the cell locating branch predicts the mask of cancerous cells. Then the masked cancerous and non-cancerous areas of the same sample are input back to the binary classification branch separately. The loopback between two branches enables the category label to supervise the cell locating branch to learn the locating ability for cancerous areas. Experiment results show that the proposed LoopNet achieves 97.5% accuracy on MVI classification. Surprisingly, the proposed loopback mechanism not only enables LoopNet to predict the cancerous area but also facilitates the classification backbone to achieve better classification performance.

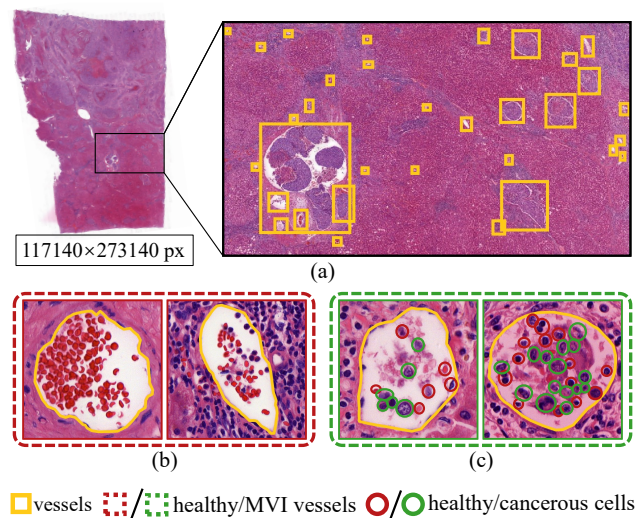


Figure 1. Examples of MVI and healthy vessels extracted from a pathological image of liver cancer. (a) The super large sample contains numerous blood vessels of varied sizes. (b) The healthy vessels are composed of a variety of cells with similar appearances. (c) The cancerous cells have varied types and similar appearances to parts of healthy cells.

1. Introduction

Microvascular invasion (MVI), referring to the appearance of cancerous cells within microscopic venules or veins, is a histological feature of cancer-related to aggressive biological behavior [27, 56]. In clinical, MVI is usually used as a reference standard for assessing cancer spreading, which is a critical factor for prognosis evaluation and treatment [8, 15, 43]. Accurate prognosis evaluation along with appropriate treatment can effectively improve patient's life quality and prolong their life-span.

Currently, the diagnosis of MVI relies on pathologists to manually find out cancerous cells from hundreds of blood vessels, each of which usually contains dozens of cells. As shown in Fig.1, each pathological sample is an image of

*Corresponding author

about $100,000 \times 250,000$ px. These super-large pathological images have three characteristics. Firstly, each sample contains numerous blood vessels (Fig.1a). Secondly, each blood vessel usually has a variety of cells with similar appearances (Fig.1b). Thirdly, types of cancerous cells are also varied (Fig.1c). Therefore, diagnosis of MVI requires the professional pathologist to discriminate cancerous/non-cancerous cells carefully, which is time-consuming and tedious. The discrimination relies on the individual pathologist's prior knowledge, which is subjective and leads to misdiagnosis occasionally.

In recent years, deep learning has achieved promising results in many areas [28–30, 68–72], including medical image analysis. Many researchers focus on applying deep learning techniques to image-based tumor analysis tasks, such as tumor grading [3, 73], lesion area detection [14, 35], vessel segmentation [18, 31], cell detection/segmentation [42, 54, 67, 75], *etc.* The successful application of deep learning relies on massive annotated samples. However, annotating cancerous cells of all MVI images is very time-consuming. What's more, the black-box characteristic of deep learning leads to unexplainable classification results, which limits the clinical application of deep learning based diagnostic methods.

In order to apply the deep learning technique to the MVI analysis task, we collect the first Pathologic Vessel Image Dataset (PVID) containing healthy blood vessel samples and MVI samples from the pathological image of liver cancer patients.

In this paper, we aim to develop an accurate, objective, and explainable method for MVI diagnosis with as few annotations as possible. As annotating the cell in each MVI vessel is time-consuming, we only adopt easily-obtained image-level category labels for developing the new approach.

For the explainable MVI classification, the developed approach should provide credible evidence, such as cancerous areas and classification results. Therefore, the proposed approach is devised to be composed of two branches: the binary classification branch and the cell locating branch. The binary classification branch is used to classify the healthy blood vessels and MVI vessels with corresponding vessel image-level category labels as supervision. The initial goal of the cell locating branch is to distinguish the cancerous cells. However, the supervision information for the cell locating branch is insufficient, which requires exploring more supervision information from the characteristic of MVI itself.

Firstly, based on the characteristic of blood vessel samples that most cells can be distributed into some similar templates according to structure and color, the correlation filter [9, 22], which is widely adopted in the object tracking area, can be used for locating most of the cells; hence the results of this filter can be interpreted as pseudo masks of cells for supervising the cell locating branch to distinguish

cell area from the background. Secondly, the healthy vessel sample only containing non-cancerous cells and background is used for supervising the cell locating branch distinguishing healthy area (non-cancerous cells and background) from the cancerous cells. Lastly, we devise loopback strategy between the binary classification branch and cell locating branch to discover the cancerous area from each MVI sample.

For the loopback strategy, the cell locating branch first predicts the cancerous area of the MVI sample, then the cancerous and non-cancerous areas of the same sample masked with the predicted results are input back into the classification branch separately. The devised a loopback strategy effectively achieves two goals: 1) utilizing the image-level category label to supervise the cell locating branch distinguishing the cancerous area from other areas. 2) building the direct relation between the predicted cancerous areas and the final classification result.

Experiment results show that the loopback strategy not only enables the proposed framework to predict precious cancerous areas but also facilitate the classification branch achieve better classification performance. The two-branch framework with the loopback strategy, termed as Loopback Network (LoopNet), achieves 97.5% accuracy on MVI classification.

In conclusion, the main contributions of our work are summarized as follows:

- We propose the first deep learning based network, termed as LoopNet, for explainable MVI classification. LoopNet fully exploits the characteristics of MVI samples to achieve blood vessel classification and cell locating results simultaneously and can be extended to MVI analysis tasks on various organs.
- The loopback strategy is devised for utilizing the category label to supervise LoopNet distinguishing the cancerous area from other regions, which effectively builds the direct relation between the located cancerous area and the final classification result.
- We collect the first Pathologic Vessel Image Dataset (PVID) containing 4130 healthy blood vessel samples and 857 MVI samples from the pathological image of 103 liver cancer patients.
- Experiment show that LoopNet achieves 97.5% accuracy on PVID, which verifies the potential of deep learning on MVI classification task.

2. Related Work

To our knowledge, there has been no MVI classification method for pathological images until now. Therefore, from a technical point of view, we survey two most related areas: explainable classification, cell detection, and segmentation.

2.1. Explainable Classification

Deep learning based techniques have been widely used in medical image classification tasks [3, 73]. However, the black box characteristic of deep learning techniques has limited their applications in clinical scenarios. The root reason is that the prediction is unexplainable. Therefore, some researchers developed some explainable classification methods, which can be broadly summarized into two categories: prediction approximation techniques and prediction attribution techniques.

For prediction approximation techniques, most methods adopted explainable machine learning methods to approximate the prediction of deep models. [7, 36] adopted explainable random forests to approximate CNN-based classification model's predictions. Further, Chen *et al.* [11] incorporated the tree-decision mechanism into the CNN-based classification model, which combines the explainability of random forests and the high performance of the CNN model. However, those methods are unsuitable for the MVI classification task, which requires knowing the number and location of cancerous cells.

For prediction attribution techniques, some feature attribution strategies are proposed for locating the critical features for final prediction results. The most commonly used strategies are activation-based, perturbation-based, and back-attribution-based techniques. The activation-based techniques [13, 44, 61, 76] attributed important features by calculating a group of weights and then summing the feature map. The perturbation-based techniques [37, 74, 77, 80] attributed important features of the input image by removing, masking, or altering them and running a forward pass on the modified image, measuring the difference with the actual output. For the back-attribution-based techniques, some researchers applied the derivative-related terms of the predict category w.r.t. the input to locate the important features. The existing derivative-related terms, including Gradient [6, 52], Gradient \times Input [5, 16, 51], Integrated Gradients [57], and DeepLIFT [50], have been proven to be firmly related or approximate by Ancona *et al.* [2] from theoretical and practical perspectives. Those feature attribution strategies built the association between important features and the final prediction. However, different methods usually attribute to different feature areas, which shows that the established association lacks credibility.

2.2. Cell Detection and Segmentation

In recent years, plenty of fully supervised methods have been proposed for cell detection [20, 54, 63, 64, 67, 79] and segmentation [1, 12, 19, 34, 42, 48, 49, 58, 60, 75]. Most of those methods leveraged manually annotated centroids/outlines/masks of cells to supervise the training of the model. Apart from the above end-to-end training methods, some researchers took the special characteristics of cells

into consideration for better detection and segmentation results. Based on the initial segmentation results with FCN, Naylor *et al.* [46] applied the watershed method to split the cells. Similarly, Xing *et al.* [65] performed bottom-up shape deformation and top-down shape inference with the initial segmentation results to achieve better cell segmentation alternately. Sirinukunwattana *et al.* [55] added the local neighborhood constraint into the cell detection and classification model. Naylor *et al.* [45] devised a regression network for cell distance map segmentation with a fully convolutional network. However, the performance of those methods highly relies on a large number of fine annotations.

For the weakly supervised methods, Xu *et al.* [66] fine-tuned a Stacked Sparse Autoencoder pretrained with image reconstruction by classifying each cell patch for detecting cells automatically. Mahmood *et al.* [41] adopted the condition GAN to segment cells with some synthetic samples and original annotated samples. LIRNet [78] adopted cascaded truncated counting indicators on image patches rather than centroid annotations to train a cell detection network. Chamanzar *et al.* [10] adopted the Voronoi transformation to generate local polygon regions containing only one cell based on centroid annotations, then trained the segmentation network with the generated pseudo annotations. Hu *et al.* [26] utilized a Generative Adversarial Network (GAN) to generate a cell centroid likelihood map, then used guided backpropagation to visualize the pixel contributions of the map, and finally obtained instance segmentation of cells by graph-cut. Feng *et al.* [17] proposed a mutual-complementing framework for detecting and segmenting cells simultaneously, where detection and segmentation branches are optimized iteratively. Most above methods are used for segmenting or detecting all cells in the pathological tissue area. Due to the unique characteristic of MVI samples, existing methods can't be directly used to detect and segment the cancerous cells. What's more, most of the above methods still require cell-level annotations.

Some researchers developed unsupervised methods to detect and segment cells to relieve the massive cost of annotations on cells. Le *et al.* [24] proposed an unsupervised cross-wise sparse convolutional autoencoder to detect cells based on the local sparsity assumption. Hou *et al.* [23] adopted GAN to synthesize histopathology samples and then trained a task-specific cell segmentation network with the synthetic samples and corresponding masks. However, the unsupervised methods usually fail in real scenarios, especially in cancerous cases, which is unsuitable for the MVI classification task.

3. Method

In order to achieve an accurate, objective and explainable diagnosis and analysis of MVI with as few annotations as possible, we devise a Loopback Network consisting of two

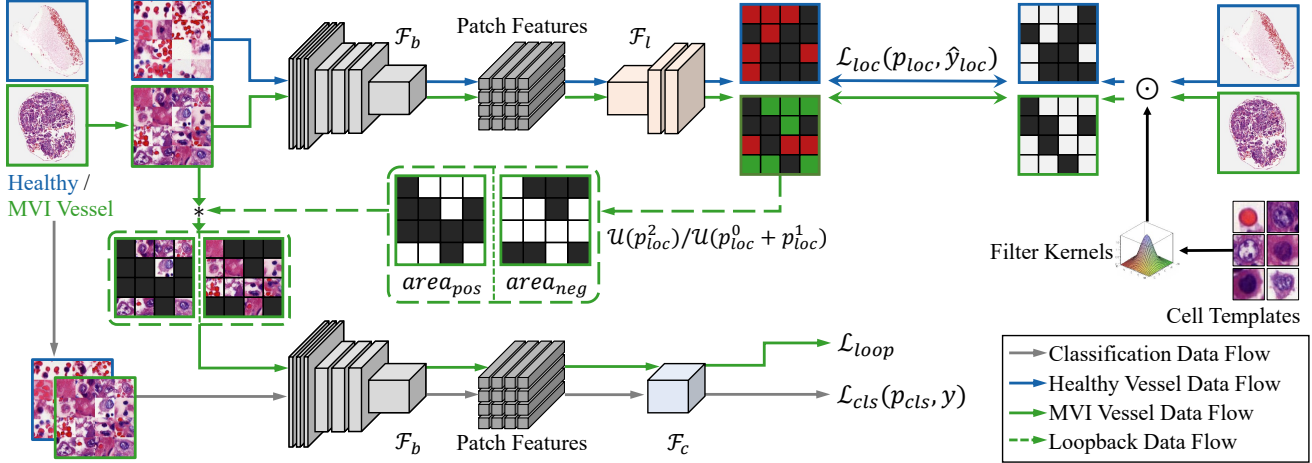


Figure 2. The framework of LoopNet, which is composed of a binary blood vessel image classification branch $\mathcal{F}_b \circ \mathcal{F}_c$ and a cell locating branch $\mathcal{F}_b \circ \mathcal{F}_l$. The image classification branch and cell locating branch share the same backbone \mathcal{F}_b , which extracts the same pathological features for classification and locating. Both healthy and MVI samples are sent to the image classification branch to predict their categories, supervised by image-level labels y annotated by the pathologist. Manually selecting some typical cells as templates, the kernelized correlation filter is used to generate pseudo mask \hat{y}_{loc} to supervise grid patch based cell locating for both kinds of vessels. For locating cancerous cells in MVI vessels, we devise a loopback strategy, which separately inputs the masked cancerous areas and masked healthy areas of an MVI vessel sample predicted by the cell locating branch into the image classification branch. The category label will ensure that the masked cancerous areas contain cancerous features and the masked healthy areas only contain healthy features. Therefore, the loopback strategy built the direct association between the locating cancerous areas and final classification results.

branches, as shown in Fig.2. One of the two branches is responsible for binary vessel image classification, which leverages the easily obtained image-level category labels as supervision.

As an indicator of MVI, the presence of cancerous cells in the blood vessel offers credible evidence for the image classification result. Therefore, we add a cell locating branch to discover cancerous cells in these blood vessels. As there are also non-cancerous cells and tissues in MVI vessels, this branch recognize three categories, including background areas, healthy cell areas and cancerous cell areas.

With only the image-level category labels annotated by the pathologist, we devise a loopback strategy to supervise locating cancerous cell areas. For MVI vessels, the healthy and cancerous areas masked with the results predicted by the cell locating branch are **input back into** the vessel image classification branch to utilize the image-level category label to supervise the cell locating branch distinguishing the cancerous area from other areas. For healthy vessels, the pseudo masks generated by the cell template correlation filter will prompt the cell locating branch to learn to recognize background and healthy cell areas. Therefore, only when the background, healthy cell areas, and cancerous cell areas of MVI vessels are perfectly distinguished will all constraints in our proposed LoopNet be satisfied. In this way, the loopback strategy can build the direct relation between the located cancerous areas and the final classification result.

The image classification branch and cell locating branch share the same backbone parameters to leverage extracted pathological image features together, followed by an image

classification head and a cell locating head, as shown in Fig.2. We denote the backbone, image classification head, and cell locating head as \mathcal{F}_b , \mathcal{F}_c , and \mathcal{F}_l , respectively. For convenience, we denote the composite function $\mathcal{F}_b \circ \mathcal{F}_c$ as \mathcal{F}_{cls} and $\mathcal{F}_b \circ \mathcal{F}_l$ as \mathcal{F}_{loc} , which correspond to the vessel image classification branch and cell locating branch, respectively.

For preprocessing, we utilize vessel segmentation [18] results to remove tissues from the vessels and only focus on the contents inside the vessel lumens. To tackle the variety of vessel sizes, we split the entire vessel image into square patches with the 3x size of the average cell size on the statistical discovery that big vessels generally contain a lot of white backgrounds. The patches containing cells determined by pseudo masks are randomly concatenated to a 20×20 patch image to obtain the fixed size input for training the network, discarding the patches that don't contain any cells. If there are less than 20×20 extracted patches, blank patches will fill the vacancy.

Based on the gained patches, each patch will contain 1-2 cells, so the patch classification can be used for cell locating, reducing the difficulty of finely segmenting the cell edge and meanwhile satisfying the requirement of analysis and diagnosis of MVI.

3.1. Binary Vessel Image Classification

For a blood vessel image I , we apply some preprocessings and data augmentations to I , and the result, termed as x , is input into the binary vessel image classification branch \mathcal{F}_{cls} to predict the probability of I belonging to an MVI vessel, denoted as $p_{cls} = \mathcal{F}_{cls}(x)$. The image-level category label

y is used to supervise the binary image classification branch with the following GHM-C loss function [38]:

$$\mathcal{L}_{cls}(p_{cls}, y) = \frac{\mathcal{L}_{CE}(p_{cls}, y)}{GD(|p_{cls} - y|)}, \quad (1)$$

where \mathcal{L}_{CE} is the cross entropy (CE) loss function, GD is gradient density function.

The reason why we adopt the GHM-C loss function [38] rather than the traditional cross-entropy loss function is that the diagnosis of MVI is a subjective task, so a few image category labels of hard blood vessel samples may be wrong annotated. The GHM-C loss function [38] can reduce the gradient contribution of massive easy examples and few outliers to make the classification network more robust.

3.2. Grid Patch Based Cell Locating

The image-level blood vessel category label, which is the only supervision information for training our model, is insufficient for locating cells in vessels. Consequently, exploring more supervision information from the characteristic of the MVI pathological image itself is necessary.

Based on the fact that most cells possess similar shapes, colors, and structures in blood vessel pathological images, distinguishing from background tissue, we can distribute the cells into some templates according to these characteristics. To distinguish background tissue and cells, the kernelized correlation filter [9, 22], as a mature technology in the object tracking field, is adopted for locating most cells in vessels with the manually selected cell templates, as shown in Fig. 2. The result of the kernelized correlation filter can be used as a binary pseudo mask for locating cells.

Due to the indistinguishable appearance of non-cancerous cells and cancerous cells, the correlation filter cannot differentiate cancerous cells from non-cancerous cells. The fact that healthy vessels only contain healthy cells can be utilized for discriminating healthy areas (non-cancerous cells and background tissue) from cancerous cells.

The presence of cancerous cells in blood vessels is the key characteristic of MVI, regardless of the positions and number of cancerous cells in vessels. Therefore we devise a loopback strategy between the image classification branch and cell locating branch to distinguish the cancerous area from the healthy area. The loopback strategy separately inputs the cancerous areas and healthy areas of an MVI vessel sample predicted by the cell locating branch into the binary image classification branch to supervise the locating of cancerous cells using image-level labels. What's more, the loopback strategy can build a direct relationship between the predicted cancerous areas and the final classification result.

3.2.1 Correlation Filter based Pseudo Mask for Distinguishing Cells and Backgrounds

Owing to the similarity of cell appearance, the correlation filter based on manually selected cell templates can locate most of the cells, which can be regarded as pseudo masks to train the cell locating branch. Specifically, splitting the correlation filter result into grid patches corresponding to the input image, every patch containing any parts of cells is labeled as positive, and patches containing only background tissue are labeled as negative. The binary pseudo mask is denoted as \hat{y}_{loc} . For an input x , the cell locating result p_{loc} for all grid patches in x is predicted by the cell locating branch \mathcal{F}_{loc} , namely $p_{loc} = \mathcal{F}_{loc}(x)$, which has three output channels representing the background, non-cancerous and cancerous areas, denoted as $p_{loc}^{0,i,j}$, $p_{loc}^{1,i,j}$, and $p_{loc}^{2,i,j}$ for a patch in row i and column j , respectively.

Since healthy blood vessels only consist of background and non-cancerous cells, the pseudo binary masks are enough to supervise cell locating with cross-entropy loss. But for MVI blood vessels consisting of cancerous and non-cancerous cells, the pseudo binary masks are insufficient for locating cells precisely. Hence, we only utilize the background areas of pseudo binary masks to supervise cell locating of MVI samples. The formula of loss function for pseudo mask cell locating is as follows:

$$\begin{aligned} \mathcal{L}_{loc} = & -\frac{\mathbf{1}\{y=0\}}{N} \sum_{i,j} \left((1 - \hat{y}_{loc}^{i,j}) \log(p_{loc}^{0,i,j}) + \hat{y}_{loc}^{i,j} \log(p_{loc}^{1,i,j}) \right) \\ & - \frac{\mathbf{1}\{y=1\}}{\sum_{i,j} \mathbf{1}\{\hat{y}_{loc}^{i,j}=0\}} \sum_{i,j} \mathbf{1}\{\hat{y}_{loc}^{i,j}=0\} \log(p_{loc}^{0,i,j}), \end{aligned} \quad (2)$$

where $N = h \times w$ is the total number of grid patches of x , h and w are number of rows and columns of grid patches, $i \in [1, h]$ and $j \in [1, w]$ is the row index and column index, $\mathbf{1}\{\cdot\}$ is the indicator function.

The \mathcal{L}_{loc} can supervise the cell locating branch learning the features of background patches and non-cancerous cell patches. But the cell locating branch still can not recognize cancerous cell patches in MVI samples, that is one of the reasons why we develop the following loopback strategy.

3.2.2 Loopback Strategy for Cancerous Cells Locating

Leveraging the characteristic that MVI blood vessels must contain one or more cancerous cells, no matter what other areas in blood vessels look like, we devise our loopback strategy to supervise locating cancerous cell patches with only image-level category labels. Specifically, for an MVI vessel sample, the cancerous areas of it can be represented by the 2-th channel of the result predicted by the cell locating branch, namely $area_{pos} = \mathcal{U}(p_{loc}^2)$, and the healthy areas of it can be represented by the sum of the 0-th channel and the 1-

th channel of the locating result, namely $area_{neg} = \mathcal{U}(p_{loc}^0 + p_{loc}^1)$. The above \mathcal{U} is the nearest neighbor interpolation function, which upsamples the patch-wise result to the size of input x .

We perform element-wise product between the original input vessel image and corresponding cancerous/healthy areas. The result $x * area_{pos}$ and $x * area_{neg}$ are **input back into** the binary image classification branch \mathcal{F}_{cls} to obtain the loopback image classification result. Ideally, the healthy areas of the MVI sample x don't contain cancerous cells, so that the loopback image classification result $\mathcal{F}_{cls}(x * area_{neg})$ will be negative. Similarly, the loopback image classification result of the MVI sample's cancerous areas $\mathcal{F}_{cls}(x * area_{pos})$ will be positive. However, the inadequately trained cell locating branch can not distinguish cancerous areas from healthy areas accurately, *e.g.* healthy areas that contain some cancerous cells will be predicted as positive, which is different from the ideal case. Therefore, we can utilize the divergence between the loopback image classification result and the ideal case to generate gradients with respect to the cell locating result p_{loc} to modify it. The formalized loss function of the loopback strategy is given as follows:

$$\begin{aligned} \mathcal{L}_{loop} = & \mathcal{L}_{CE}\left(\mathcal{F}_{cls}\left(x * \mathcal{U}(p_{loc}^0 + p_{loc}^1)\right), 0\right) \\ & + \mathcal{L}_{CE}\left(\mathcal{F}_{cls}\left(x * \mathcal{U}(p_{loc}^2)\right), 1\right). \end{aligned} \quad (3)$$

This loss function is only applied to MVI samples.

It is worth noting that the gradients of \mathcal{L}_{loop} w.r.t. the parameters of image classification branch \mathcal{F}_{cls} in Eq.(3) are not accumulated to participate in gradient descent. Since in loopback procedure, the role of the image classification branch is like the discriminator in the generative adversarial network, optimizing the parameters of \mathcal{F}_{cls} will make the discrimination ability degenerate.

The loopback strategy \mathcal{L}_{loop} can supervise LoopNet's cell locating branch to learn the cancerous cell patches' features through the gradients of \mathcal{L}_{loop} back-propagating to \mathcal{F}_{loc} . Moreover, the loopback strategy correlates the vessel image classification results with the cancerous cell locating results, which provides credible evidence and a reliable explanation of the image classification result and facilitates the classification branch to achieve better classification performance.

3.3. Complete Algorithm

Based on the well-developed image classification network, the loopback strategy, together with the binary pseudo mask, enables our LoopNet to obtain the ability of distinguishing healthy and MVI blood vessels, and the capability of locating background, non-cancerous cells and cancerous cells in vessels, supervised only by the image-level category labels. LoopNet learns the features of background and non-cancerous cell patches through pseudo binary masks

generated by cell template correlation filter and learns the features of cancerous cell patches through through loopback strategy performed on MVI samples.

We train our LoopNet in two stages. Firstly, we train the binary vessel image classification branch \mathcal{F}_{cls} using the loss function \mathcal{L}_{cls} to endow it with the essential ability to distinguish healthy/MVI blood vessels. Secondly, we jointly optimize \mathcal{F}_{cls} and the cell locating branch \mathcal{F}_{loc} with the following loss function:

$$\mathcal{L} = \mathcal{L}_{loc} + \alpha \mathcal{L}_{loop} + \beta \mathcal{L}_{cls}, \quad (4)$$

where α and β are balance parameters.

4. Dataset

To construct the Pathologic Vessel Image Dataset (PVID) for MVI analysis, we collect 100 whole slide pathological images of liver cancer patients from the cooperative institution, each of which is about $100,000 \times 250,000$ pixels. For each whole slide pathological image, we randomly crop an average of about 50 blood vessel images, which can reduce the repeatability of the sample. Finally, the collected PVID contains 5,000 vessel samples (4,140 healthy blood vessels and 860 MVI blood vessels). We randomly split the vessel image part into the training, validation, and test sets according to slides, ensuring vessel images of the same slide are in the same set. The number of samples of healthy vessels in the training set, validation set and test set is 2,480 and 830 and 830 respectively, and the number of MVI vessels is 520 and 170 and 170 respectively.

For both vessel image classification and cell locating, we only use the training set of the vessel image part of PVID to train the proposed model. The test set of the vessel image part is adopted to evaluate the performance of vessel image classification.

To assess the cancerous cell locating performance, we select another 130 MVI blood vessels and annotate the centroids of every cancerous cell in these blood vessels using point label as the locating test part of PVID. The number of annotated cancerous cell centroids is 23,237.

All the annotations were labeled by the experienced pathologist of the cooperative institution. What's more, this study was approved by the institutional research ethics committee.

5. Experiments

5.1. Network Architecture and Parameters

In the following experiment, unless otherwise specified, the backbone \mathcal{F}_b we adopted is the ResNet-50 [21], and the last global average pooling layer and the fully connected layer are removed and regarded as vessel image classification head \mathcal{F}_c . The cell locating head we adopted is a 3-

Index\Backbone		ResNet-50 [21]		AlexNet [33]		VGG-16 [53]		EfficientNetV2-S [59]		MobileNetV2 [25]		ConvNeXt-B [40]	
		Baseline	Ours	Baseline	Ours	Baseline	Ours	Baseline	Ours	Baseline	Ours	Baseline	Ours
Accuracy		96.59	97.49	96.09	96.19	97.09	97.99	96.89	97.59	94.68	97.19	94.58	95.29
Precision	Healthy Vessel	97.71	98.55	96.90	98.52	97.27	98.79	97.83	98.20	94.58	97.61	97.67	98.61
	MVI Vessel	91.02	92.44	91.77	85.95	94.15	96.10	92.17	94.55	95.38	94.97	78.82	84.44
Recall	Healthy Vessel	98.18	98.43	96.85	98.32	99.27	98.79	98.43	98.91	99.03	99.27	94.79	96.61
	MVI Vessel	88.89	92.98	84.80	92.98	86.55	94.15	89.47	91.23	72.51	88.30	93.57	88.89

Table 1. The classification performance comparison with SOTA methods. The *Baseline* means directly using the backbone for image classification. *Our* LoopNet shares the same backbone with the corresponding SOTA classification networks. (All scores are in %)

Index\Method	CAM [44]	DeepLIFT [50]	LRP [4]	LoopNet
Precision	4.59	81.11	39.05	71.26
Recall	8.35	5.39	5.89	94.52
Dice	16.35	15.09	15.42	79.58

Table 2. The comparison results of classification explainability. The precision, recall, and dice score denote the performance of the located cancerous patch in the attributed feature areas for CAM [44], DeepLIFT [50] and LRP [4], which are three classic feature attribution methods.

layers Graph Convolutional Network (GCN) [32]. Batch-Norm, ReLU, and Dropout layers are successively added between two GCN layers. The dropout probability of the first and second Dropout layer is set to 0.2 and 0.1, respectively. Edges are added between each feature and its eight spatial neighbors.

We solely train the image classification branch 50 epochs and then train the entire LoopNet 50 epochs with $\alpha = 1$ and $\beta = 0.5$. The default batch size is 16. The optimizer we adopted is Ranger [62], and the learning rate and weight decay we set are 1×10^{-3} and 5×10^{-4} , respectively. The cosine annealing with five warm-up epochs is adopted as the learning rate scheduler. The size of the input image is 640×640 px.

5.2. Performance of MVI Classification

As described above, there is no MVI classification method until now. Therefore, we compare the proposed LoopNet with the SOTA fully supervised CNN-based image classification methods, including ResNet-50 [21], AlexNet [33], VGG-16 [53], EfficientNetV2-S [59], MobileNetV2 [25], and ConvNeXt-B [40]. For a fair comparison, we adopt the same backbone as those classification models.

The vessel classification performance comparison with the SOTA methods is given in Table 1, where we can see that the proposed LoopNet achieves the best accuracy, precision, and recall in almost all cases than SOTA classification methods. This is mainly because the proposed loopback strategy builds direct relation between the predicted cancerous areas and the image classification result, prompting the network to pay more attention to the discriminative cancerous areas rather than irrelevant background areas.

Index\Method	Qu [47]	Liu [39]	Fully	LoopNet
Precision	85.47	80.82	78.99	71.26
Recall	95.61	92.57	93.35	94.52
Dice	89.34	84.42	82.07	79.58

Table 3. The cancerous cell locating results of different methods. [47] and [39] are two cell segmentation approaches with cell centroid point as annotations. ‘Fully’ means using the cancerous cell centroid annotations to finetune the cell locating branch of LoopNet, of which the image classification branch has been trained.

5.3. Comparison of Classification Explainability

Another advantage of the proposed LoopNet is providing explainable classification results, namely the cancerous cell areas. To verify the effectiveness of explainability, we compare the cancerous cell area locating results with the outputs of three classic feature attribution methods on the test part of PVID: CAM [44], DeepLIFT [50] and LRP [4]. Those methods adopted different feature attribution strategy to locate the critical features for the final classification prediction. Table 2 shows that the proposed LoopNet achieves the best performance among all methods. Noting that the Precision of DeepLIFT is higher than our approach, the Recall and Dice of it is much lower. Besides, the all indexes of other two approaches are pretty lower than the proposed LoopNet.

The qualitative visual results of different methods for the located cancerous areas and detailed analysis are given in the *supplements*.

5.4. Performance Comparison of Cell Locating

To assess the effectiveness of the proposed loopback strategy for cell locating only with image-level category labels, we compare the cell locating performance to two weakly supervised cell segmentation approaches, Qu [47] and Liu [39], both leveraging point annotations of cells. The results are shown in Table 3. ‘Fully’ means using the cancerous cell centroid annotations to directly finetune the cell locating branch of LoopNet where the image classification branch has been trained. The results indicate that only with image-level category labels the performance of our approach can achieve promising results compared to those approaches that require massive point annotations.

Index\Ablation	w/o \mathcal{L}_{loop}	w/o \mathcal{L}_{loc}	w/o \mathcal{L}_{cls}	LoopNet
Precision	0	50.86	60.58	71.26
Recall	0	99.75	99.50	94.52
Dice	0	59.93	73.03	79.58

Table 4. The results of ablation study on different loss terms.

5.5. Ablation Study

5.5.1 Ablation of Loss Terms

As delineated in Sec.3.3, the joint training loss function has three terms: the image classification loss \mathcal{L}_{cls} , the pseudo mask based cell locating loss \mathcal{L}_{loc} and the loopback loss \mathcal{L}_{loop} . We evaluate the joint training when supervised with the three terms separately in Table 4, showing that all loss terms have contributed to the final result. We have to mention that without \mathcal{L}_{loop} , the model will degenerate so that it can't find any cancerous cells and all the metrics will be 0, demonstrating that the devised loopback strategy plays the key role in locating cancerous cells. Abandoning \mathcal{L}_{cls} or \mathcal{L}_{loc} will cause the network to recognize many healthy areas as cancerous, so the Recall becomes very high but Precision and Dice decrease a lot. The \mathcal{L}_{loc} acts as an antagonist and constraint to \mathcal{L}_{loop} , and the \mathcal{L}_{cls} can maintain the discriminative capability of the image classification branch for optimizing the cell locating results with the proposed loopback strategy. Accordingly, the three terms work together to improve cancerous cell locating performance.

5.5.2 Ablation of Influence of GHM-C Loss

As described in Sec.3.1, owing to the subjectivity of MVI diagnosis, there will be a few incorrectly labeled vessel samples. Therefore we adopt GHM-C loss rather than CE loss, which can reduce the gradient contribution of these incorrectly labeled outliers. As shown in the second column of Table 5, compared to the original result in the first column, replacing the GHM-C loss of \mathcal{L}_{cls} to CE loss will reduce the overall Accuracy and especially the Recall of MVI vessels, which is sensitivity, an important index in clinical.

In \mathcal{L}_{loop} , we adopt CE loss for modifying cancerous cell locating results using image-level labels rather than GHM-C loss. That is because the masked cancerous or non-cancerous areas of the original input vessel image can be regarded as the hard sample for the image classification branch, requiring more gradients magnitude for modifying the cancerous cell locating results. As shown in the third column in Table 5, replacing the CE loss in \mathcal{L}_{loop} with GHM-C loss will induce the network to recognize more cancerous areas, making the masked samples more like easy samples, therefore reducing the Precision and Dice.

6. Conclusion

In this paper, we put forward the first deep learning based network LoopNet for classifying MVI, which can be used

Index\Ablation		Ori	\mathcal{L}_{cls} GHM-C → CE	\mathcal{L}_{loop} CE → GHM-C	
Vessel Classification	Accuracy	97.49	96.59	-	
	Precision	<i>Healthy</i>	98.55	97.37	-
		<i>MVI</i>	92.44	92.13	-
	Recall	<i>Healthy</i>	98.43	98.67	-
		<i>MVI</i>	92.98	87.13	-
	Cell locating	Precision	71.26	-	63.26
Recall		94.52	-	98.95	
Dice		79.58	-	75.97	

Table 5. The ablation of replacing the GHM-C loss and CE loss in \mathcal{L}_{cls} and \mathcal{L}_{loop} to the other. 'Ori' denotes the original setting.

as an accurate, objective, explainable and efficient diagnosis tool for MVI. Through fully exploiting the characteristic of MVI samples, LoopNet can achieve blood vessel classification and cell locating results simultaneously with only category labels, which provides a new weakly supervised framework for future MVI analysis tasks on various organs. To achieve the explainable MVI analysis goal, the loopback strategy is devised for utilizing the category label to supervise LoopNet distinguishing the cancerous area from other regions, which effectively builds the direct relation between the located cancerous area and the final classification result. To verify the effectiveness of the proposed the LoopNet, we collect the first Pathologic Vessel Image Dataset (PVID). Experiment results demonstrate that the proposed LoopNet achieves 97.5% accuracy on PVID, which demonstrates the potential of deep learning on the MVI analysis task.

The cell locating results show that there are still several missing cancerous cells, which demonstrates the deficiency of insufficient supervised information. In the future, we will focus on exploring more potential supervision information from the characteristics of pathologic images and incorporating those supervision information into the proposed framework. Furthermore, we will also devote ourselves to improving the overall performance of the proposed method and applying the proposed method to auxiliary diagnosis in clinical practice.

Acknowledgements. This work is supported by National Natural Science Foundation of China (61976186,U20B2066), Zhejiang Provincial Science and Technology Project for Public Welfare (LGF21F020020), Starry Night Science Fund of Zhejiang University Shanghai Institute for Advanced Study (Grant No. SN-ZJU-SIAS-001), Fundamental Research Funds for the Central Universities (2021FZZX001-23), Alibaba Group through Alibaba Innovative Research Program, and Alibaba-Zhejiang University Joint Research Institute of Frontier Technologies.

References

- [1] Hao Chen A, Xiaojuan Qi A, Lequan Yu A, Qi Dou A, Jing Qin B, and Pheng Ann Heng A. Dcan: Deep contour-aware networks for object instance segmentation from histology images - sciencedirect. *Medical Image Analysis*, 36:135–146, 2017. 3
- [2] Marco Ancona, Enea Ceolini, Cengiz Zireli, and Marko Gross. Towards better understanding of gradient-based attribution methods for deep neural networks. In *ICLR*, 2018. 3
- [3] Eirini Arvaniti, Kim S Fricker, Michael Moret, Niels Rupp, Thomas Hermanns, Christian Fankhauser, Norbert Wey, Peter J Wild, Jan H Rueschoff, and Manfred Claassen. Automated gleason grading of prostate cancer tissue microarrays via deep learning. *Scientific reports*, 8(1):1–11, 2018. 2, 3
- [4] S. Bach, A. Binder, G. Montavon, F. Klauschen, KR Müller, and W. Samek. On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PLOS ONE*, 10, 2015. 7
- [5] Sebastian Bach, Alexander Binder, Grégoire Montavon, Frederick Klauschen, Klaus Robert Müller, and Wojciech Samek. On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PLOS ONE*, 10(7):130140, 2015. 3
- [6] David Baehrens, Timon Schroeter, Stefan Harmeling, Motoaki Kawanabe, Katja Hansen, and Klaus-Robert Müller. How to explain individual classification decisions. *Journal of Machine Learning Research*, 11(61):1803–1831, 2010. 3
- [7] Osbert Bastani, Carolyn Kim, and Hamsa Bastani. Interpretability via model extraction. *arXiv preprint arXiv:1706.09773*, 2017. 3
- [8] Rubén G. Bengió, Leandro Cristian Arribillaga, Javier Epelde, Sergio Orellana, Ariel Montedoro, Verónica Bengió, Esteban Cordero, and Matías Guevara. Evaluation of microvascular invasion as a prognostic factor in the progression of non-metastatic renal cancer. *Central European Journal of Urology*, 71(4), 2018. 1
- [9] David S Bolme, J Ross Beveridge, Bruce A Draper, and Yui Man Lui. Visual object tracking using adaptive correlation filters. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 2544–2550. IEEE, 2010. 2, 5
- [10] Alireza Chamanzar and Yao Nie. Weakly supervised multi-task learning for cell detection and segmentation. In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pages 513–516. IEEE, 2020. 3
- [11] Ying Chen, Feng Mao, Jie Song, Xinchao Wang, Huiqiong Wang, and Mingli Song. Self-born wiring for neural trees. *CVPR*, 2021. 3
- [12] Yuxin Cui, Guiying Zhang, Zhonghao Liu, Zheng Xiong, and Jianjun Hu. A deep learning algorithm for one-step contour aware nuclei segmentation of histopathological images. *Medical Biological Engineering Computing*, 2018. 3
- [13] Saurabh Desai and Harish G. Ramaswamy. Ablation-cam: Visual explanations for deep convolutional network via gradient-free localization. In *WACV*, pages 983–991, 2020. 3
- [14] Neeraj Dhungel, Gustavo Carneiro, and Andrew P Bradley. Deep learning and structured prediction for the segmentation of mass in mammograms. In *International Conference on Medical image computing and computer-assisted intervention*, pages 605–612. Springer, 2015. 2
- [15] S. Feng, X. Yu, W. Liang, X. Li, W. Zhong, W. Hu, H. Zhang, Z. Feng, M. Song, and J. Zhang. Development of a deep learning model to assist with diagnosis of hepatocellular carcinoma. 2021. 1
- [16] Yao Feng, Fan Wu, Xiaohu Shao, Yanfeng Wang, and Xi Zhou. Joint 3d face reconstruction and dense alignment with position map regression network. In *ECCV*, pages 557–574, 2018. 3
- [17] Zunlei Feng, Zhonghua Wang, Xinchao Wang, Yining Mao, Thomas Li, Jie Lei, Yuexuan Wang, and Mingli Song. Mutual-complementing framework for nuclei detection and segmentation in pathology image. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4036–4045, 2021. 3
- [18] Zunlei Feng, Zhonghua Wang, Xinchao Wang, Xiuming Zhang, Lechao Cheng, Jie Lei, Yuexuan Wang, and Mingli Song. Edge-competing pathological liver vessel segmentation with limited labels. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 1325–1333, 2021. 2, 4
- [19] Simon Graham, Quoc Dang Vu, Shan E Ahmed Raza, Ayesha Azam, Yee Wah Tsang, Jin Tae Kwak, and Nasir Rajpoot. Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Medical Image Analysis*, 58:101563, 2019. 3
- [20] Yue Guo, Jason Stein, Guorong Wu, and Ashok Krishnamurthy. Sau-net: A universal deep network for cell counting. In *Proceedings of the 10th ACM international conference on bioinformatics, computational biology and health informatics*, pages 299–306, 2019. 3
- [21] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016. 6, 7
- [22] João F Henriques, Rui Caseiro, Pedro Martins, and Jorge Batista. High-speed tracking with kernelized correlation filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(3):583–596, 2014. 2, 5
- [23] Le Hou, Ayush Agarwal, Dimitris Samaras, Tahsin M. Kurc, and Joel H. Saltz. Robust histopathology image analysis: To label or to synthesize? In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019. 3
- [24] Le Hou, Vu Nguyen, Ariel B. Kanevsky, Dimitris Samaras, Tahsin M. Kurc, Tianhao Zhao, Rajarsi R. Gupta, Yi Gao, Wenjin Chen, and David and Foran. Sparse autoencoder for unsupervised nucleus detection and representation in histopathology images. *Pattern recognition*, 2019. 3
- [25] Andrew G Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*, 2017. 7

- [26] Wei Hu, Huanhuan Sheng, Jing Wu, Yining Li, Tianyi Liu, Yonghao Wang, and Yuan Wen. Generative adversarial training for weakly supervised nuclei instance segmentation. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 3649–3654. IEEE, 2020. 3
- [27] Hai Huang, Xiu-Wu Pan, Yi Huang, Dan-Feng Xu, Xin-Gang Cui, Lin Li, Yi Hong, Lu Chen, Yi Gao, and Lei Yin. Microvascular invasion as a prognostic indicator in renal cell carcinoma: a systematic review and meta-analysis. In *International Journal of Clinical and Experimental Medicine*, 2015. 1
- [28] Yongcheng Jing, Yining Mao, Yiding Yang, Yibing Zhan, Mingli Song, Xinchao Wang, and Dacheng Tao. Learning graph neural networks for image style transfer. In *ECCV*, 2022. 2
- [29] Yongcheng Jing, Yiding Yang, Xinchao Wang, Mingli Song, and Dacheng Tao. Amalgamating knowledge from heterogeneous graph neural networks. In *CVPR*, 2021. 2
- [30] Yongcheng Jing, Yiding Yang, Xinchao Wang, Mingli Song, and Dacheng Tao. Meta-aggregator: learning to aggregate for 1-bit graph neural networks. In *ICCV*, 2021. 2
- [31] N. Kaur, G. Chetty, and L. Singh. A novel approach using deep neural network vessel segmentation & retinal disease detection. In *CSDE*, 2020. 2
- [32] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016. 7
- [33] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, 2017. 7
- [34] Neeraj Kumar, Ruchika Verma, Sanuj Sharma, Surabhi Bhargava, Abhishek Vahadane, and Amit Sethi. A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE Transactions on Medical Imaging*, pages 1–1, 2017. 3
- [35] Viksit Kumar, Jeremy M Webb, Adriana Gregory, Max Denis, Duane D Meixner, Mahdi Bayat, Dana H Whaley, Mostafa Fatemi, and Azra Alizad. Automated and real-time segmentation of suspicious breast masses using convolutional neural network. *PloS one*, 13(5):e0195816, 2018. 2
- [36] Jie Lei, Zhe Wang, Zunlei Feng, Mingli Song, and Jiajun Bu. Understanding the prediction process of deep networks by forests. In *BigMM*, pages 1–7, 2018. 3
- [37] Benjamin J. Lengerich, Sandeep Konam, Eric P. Xing, Stephanie Rosenthal, and Manuela M. Veloso. Visual explanations for convolutional neural networks via input resampling. *arXiv preprint arXiv:1707.09641*, 2017. 3
- [38] Buyu Li, Yu Liu, and Xiaogang Wang. Gradient harmonized single-stage detector. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 8577–8584, 2019. 5
- [39] Weizhen Liu, Qian He, and Xuming He. Weakly supervised nuclei segmentation via instance learning. *arXiv preprint arXiv:2202.01564*, 2022. 7
- [40] Zhuang Liu, Hanzi Mao, Chao-Yuan Wu, Christoph Feichtenhofer, Trevor Darrell, and Saining Xie. A convnet for the 2020s. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11976–11986, 2022. 7
- [41] Faisal Mahmood, Daniel Borders, Richard J. Chen, Gregory N. McKay, and Nicholas J. Durr. Deep adversarial training for multi-organ nuclei segmentation in histopathology images. *IEEE Transactions on Medical Imaging*, PP(99):1–1, 2019. 3
- [42] Faisal Mahmood, Daniel Borders, Richard J Chen, Gregory N McKay, Kevan J Salimian, Alexander Baras, and Nicholas J Durr. Deep adversarial training for multi-organ nuclei segmentation in histopathology images. *IEEE transactions on medical imaging*, 39(11):3257–3267, 2019. 2, 3
- [43] Rodríguez-Perálvarez Manuel, Luong Vinh Tu, Andreana Lorenzo, Meyer Tim, Paul Dhillon Amar, and Kenneth Burroughs Andrew. A systematic review of microvascular invasion in hepatocellular carcinoma: Diagnostic and prognostic variability. 2012. 1
- [44] Rakshit Naidu and Joy Michael. Ss-cam: Smoothed scorecam for sharper visual feature localization. *arXiv preprint arXiv:2006.14255*, 2020. 3, 7
- [45] Naylor, Peter, Lae, Marick, Reyal, Fabien, Walter, and Thomas. Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE Transactions on Medical Imaging*, 38(2):448–459, 2019. 3
- [46] Peter Naylor, Marick Laé, Fabien Reyal, and Thomas Walter. Nuclei segmentation in histopathology images using deep neural networks. In *IEEE International Symposium on Biomedical Imaging*, 2017. 3
- [47] Hui Qu, Pengxiang Wu, Qiaoying Huang, Jingru Yi, Gregory M Riedlinger, Subhajyoti De, and Dimitris N Metaxas. Weakly supervised deep nuclei segmentation using points annotation in histopathology images. In *International Conference on Medical Imaging with Deep Learning*, pages 390–400. PMLR, 2019. 7
- [48] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2015. 3
- [49] Monjoy Saha and Chandan Chakraborty. Her2net: A deep framework for semantic segmentation and classification of cell membranes and nuclei in breast cancer evaluation. *IEEE Transactions on Image Processing*, 27(5):2189–2200, 2018. 3
- [50] Avanti Shrikumar, Peyton Greenside, and Anshul Kundaje. Learning important features through propagating activation differences. In *International conference on machine learning*, pages 3145–3153. PMLR, 2017. 3, 7
- [51] Avanti Shrikumar, Peyton Greenside, Anna Shcherbina, and Anshul Kundaje. Not just a black box: Learning important features through propagating activation differences. *arXiv preprint arXiv:1605.01713*, 2016. 3
- [52] Karen Simonyan, Andrea Vedaldi, and Andrew Zisserman. Deep inside convolutional networks: Visualising image classification models and saliency maps. In *ICLR Workshop*, 2013. 3

- [53] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014. 7
- [54] Korsuk Sirinukunwattana, Shan E Ahmed Raza, Yee-Wah Tsang, David RJ Snead, Ian A Cree, and Nasir M Rajpoot. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. *IEEE transactions on medical imaging*, 35(5):1196–1206, 2016. 2, 3
- [55] Korsuk Sirinukunwattana, Shan E Ahmed Raza, Yee Wah Tsang, David R. J. Snead, Ian A. Cree, and Nasir M. Rajpoot. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. *IEEE Transactions on Medical Imaging*, 35(5):1196–1206, 2016. 3
- [56] Shuji Sumie, Ryoko Kuromatsu, Koji Okuda, Eiji Ando, Akio Takata, Nobuyoshi Fukushima, Yasutomo Watanabe, Masamichi Kojiro, and Michio Sata. Microvascular invasion in patients with hepatocellular carcinoma and its predictable clinicopathological factors. In *Annals of Surgical Oncology*, 2008. 1
- [57] Mukund Sundararajan, Ankur Taly, and Qiqi Yan. Axiomatic attribution for deep networks. In *ICML*, pages 3319–3328. PMLR, 2017. 3
- [58] Mahmood Tahir, Owais Muhammad, Noh Kyoung Jun, Yoon Hyo Sik, Haider Adnan, Sultan Haseeb, and Park Kang Ryoun. Artificial intelligence-based segmentation of nuclei in multi-organ histopathology images: Model development and validation. *JMIR Medical Informatics*, 2021. 3
- [59] Mingxing Tan and Quoc Le. Efficientnetv2: Smaller models and faster training. In *ICML*, pages 10096–10106, 2021. 7
- [60] Jeya Maria Jose Valanarasu, Poojan Oza, Ilker Hacihaliloglu, and Vishal M Patel. Medical transformer: Gated axial-attention for medical image segmentation. *arXiv preprint arXiv:2102.10662*, 2021. 3
- [61] Haofan Wang, Zifan Wang, Mengnan Du, Fan Yang, Zijian Zhang, Sirui Ding, Piotr Mardziel, and Xia Hu. Score-cam: Score-weighted visual explanations for convolutional neural networks. In *CVPR Workshops*, pages 111–119, 2020. 3
- [62] Less Wright. Ranger - a synergistic optimizer. <https://github.com/lessw2020/Ranger-Deep-Learning-Optimizer>, 2019. 7
- [63] Chensu Xie, Chad M Vanderbilt, Anne Grabenstetter, and Thomas J Fuchs. Voca: cell nuclei detection in histopathology images by vector oriented confidence accumulation. In *International Conference on Medical Imaging with Deep Learning*, pages 527–539. PMLR, 2019. 3
- [64] Yuanpu Xie, Fuyong Xing, Xiaoshuang Shi, Xiangfei Kong, Hai Su, and Lin Yang. Efficient and robust cell detection: A structured regression approach. *Medical image analysis*, 44:245–254, 2018. 3
- [65] Fuyong Xing, Yuanpu Xie, and Lin Yang. An automatic learning-based framework for robust nucleus segmentation. *IEEE Transactions on Medical Imaging*, 35(2):550–566, 2016. 3
- [66] Jun Xu, Lei Xiang, Renlong Hang, and Jianzhong Wu. Stacked sparse autoencoder (ssae) for nuclei detection on breast cancer histopathology images. In *IEEE International Symposium on Biomedical Imaging*, 2014. 3
- [67] Jun Xu, Lei Xiang, Qingshan Liu, Hannah Gilmore, Jianzhong Wu, Jinghai Tang, and Anant Madabhushi. Stacked sparse autoencoder (ssae) for nuclei detection on breast cancer histopathology images. *IEEE transactions on medical imaging*, 35(1):119–130, 2015. 2, 3
- [68] Xingyi Yang, Jingwen Ye, and Xinchao Wang. Factorizing knowledge in neural networks. In *Computer Vision—ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XXXIV*, pages 73–91. Springer, 2022. 2
- [69] Xingyi Yang, Daquan Zhou, Songhua Liu, Jingwen Ye, and Xinchao Wang. Deep model reassembly. *NeurIPS*, 2022. 2
- [70] Jingwen Ye, Yifang Fu, Jie Song, Xingyi Yang, Songhua Liu, Xin Jin, Mingli Song, and Xinchao Wang. Learning with recoverable forgetting. In *ECCV*, 2022. 2
- [71] Jingwen Ye, Yixin Ji, Xinchao Wang, Xin Gao, and Mingli Song. Data-free knowledge amalgamation via group-stack dual-gan. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 12513–12522, 2020. 2
- [72] Jingwen Ye, Yining Mao, Jie Song, Xinchao Wang, Cheng Jin, and Mingli Song. Safe distillation box. In *AAAI Conference on Artificial Intelligence*, 2021. 2
- [73] Xiaotian Yu, Zunlei Feng, Mingli Song, Yuexuan Wang, Xiuming Zhang13, and Thomas Li. Tendentious noise-rectifying framework for pathological hcc grading. In *British Machine Vision Conference*, 2021. 2, 3
- [74] Matthew D. Zeiler and Rob Fergus. Visualizing and understanding convolutional networks. In *ECCV*, 2014. 3
- [75] Donghao Zhang, Yang Song, Siqi Liu, Dagan Feng, Yue Wang, and Weidong Cai. Nuclei instance segmentation with dual contour-enhanced adversarial network. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 409–412. IEEE, 2018. 2, 3
- [76] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba. Learning deep features for discriminative localization. *IEEE Computer Society*, 2016. 3
- [77] Jian Zhou and Olga G Troyanskaya. Predicting effects of noncoding variants with deep learning-based sequence model. *Nature Methods*, 12(10):931–934, 2015. 3
- [78] Xiao Zhou, Zhen Cheng, Miao Gu, and Fei Chang. Lirnet: Local integral regression network for both strongly and weakly supervised nuclei detection. In *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 945–951. IEEE, 2020. 3
- [79] Yin Zhou, Hang Chang, Kenneth E. Barner, and Bahram Parvin. Nuclei segmentation via sparsity constrained convolutional regression. In *IEEE International Symposium on Biomedical Imaging*, 2015. 3
- [80] Luisa M Zintgraf, Taco S Cohen, Tameem Adel, and Max Welling. Visualizing deep neural network decisions: Prediction difference analysis. In *ICLR*, 2017. 3