

S³R: Self-supervised Spectral Regression for Hyperspectral Histopathology Image Classification

Xingran Xie, Yan Wang^(✉), and Qingli Li

Shanghai Key Laboratory of Multidimensional Information Processing,
East China Normal University, Shanghai 200241, China
51215904112@stu.ecnu.edu.cn, ywang@cee.ecnu.edu.cn, qlli@cs.ecnu.edu.cn

Abstract. Benefited from the rich and detailed spectral information in hyperspectral images (HSI), HSI offers great potential for a wide variety of medical applications such as computational pathology. But, the lack of adequate annotated data and the high spatio-spectral dimensions of HSIs usually make classification networks prone to overfit. Thus, learning a general representation which can be transferred to the downstream tasks is imperative. To our knowledge, no appropriate self-supervised pre-training method has been designed for histopathology HSIs. In this paper, we introduce an efficient and effective Self-supervised Spectral Regression (S³R) method, which exploits the low rank characteristic in the spectral domain of HSI. More concretely, we propose to learn a set of linear coefficients that can be used to represent one band by the remaining bands via masking out these bands. Then, the band is restored by using the learned coefficients to reweight the remaining bands. Two pre-text tasks are designed: (1) S³R-CR, which regresses the linear coefficients, so that the pre-trained model understands the inherent structures of HSIs and the pathological characteristics of different morphologies; (2) S³R-BR, which regresses the missing band, making the model to learn the holistic semantics of HSIs. Compared to prior arts *i.e.*, contrastive learning methods, which focuses on natural images, S³R converges at least 3 times faster, and achieves significant improvements up to 14% in accuracy when transferring to HSI classification tasks.

Keywords: Self-supervised learning · Hyperspectral histopathology image classification · Low-rank.

1 Introduction

Histopathology plays an important role in the diagnostic and therapeutic aspects of modern medicine [21,15]. As artificial intelligence is evolving rapidly, deep learning based image processing techniques have been extensively reported for histopathological diagnosis [15]. Nowadays, the microscopy hyperspectral imaging system has become an emerging research hotspot in the field of medical image

analysis [17], benefited from the rich spatio-spectral information provided by hyper-spectral images (HSI). Thus, it provides a new perspective for computational pathology and precision medicine.

Supervised learning for histopathology image analysis requires a large amount of annotations [16]. Since both expertise and time are needed, the labeled data with high quality annotations are usually expensive to acquire. This situation is more conspicuous for annotating HSIs [22], whose appearances are different compared with RGB images, so pathologists may take longer time in recognizing cancer tissues on HSIs. Typically, an HSI is presented as a hypercube, such high spatio-spectral dimensions make it difficult to perform accurate HSI classification, especially when the annotated data is limited, which may lead to overfitting problems. Thus, an appropriate HSI pre-training method is imperative.

In recent years, self-supervised pre-training has been successful in both natural language processing [6] and computer vision. Previous research of self-supervised learning mainly focuses on contrastive learning, *e.g.*, MOCO [12], SimCLR [3], BYOL [9] and SimSiam [4], training encoders to compare positive and negative sample pairs. Among these methods, BYOL [9] and SimSiam [4] achieve higher performance than previous contrastive methods without using negative sample pairs. But, the training setup *e.g.*, batch size in original papers is not always affordable for research institutions. More recently, masked image modeling (MIM) methods represented by MAE [11] have been proved to learn rich visual representations and significantly improve the performance on downstream tasks [18]. After randomly adding masks to input images, a pixel-level regression target is set as a pretext task. Almost all the MIM methods are designed based on transformers which receive and process tokenized inputs.

Different from natural images, HSIs analyze how light transmittance varies on different forms of tissues or cells, which infect the brightness of various regions [8]. This generates dozens even hundreds of narrow and contiguous spectral bands in the spectral dimension. Understanding the inherent spectral structure in self-supervised pre-training is non-trivial for the networks to conduct downstream tasks. To the best of our knowledge, there is not any self-supervised method designing the architecture tailored for microscopy HSI classification yet. In this work, we present **Self-Supervised Spectral Regression** (S^3R), an efficient and effective pre-training objective that takes advantage of low rankness in the spectral domain. More specifically, we assume that one band can be represented as a linear combination of the remaining bands, and propose to learn the linear coefficients by a Convolutional Neural Network (CNN) based backbone via masking out the remaining bands. We propose two alternative pretext tasks, which have different focuses. (1) **Coefficients Regression** (S^3R -CR), which makes the network to directly regress the “groundtruth” linear coefficients learned beforehand. In this way, the pre-trained model acquires an adequate understanding of complex spectral structures and pathological characteristics of different morphologies. (2) **Band Regression** (S^3R -BR), regressing the pixel intensity of the selected band by re-weighting the remaining bands. Since detailed information (edges, textures, etc.) is already stored in different bands of the HSI, our spectral

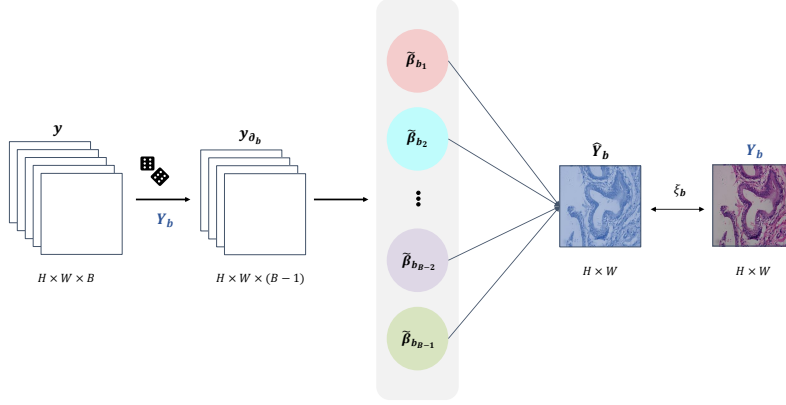


Fig. 1. Illustration of spectral regression in HSI. The whole process to estimate the coefficient $\tilde{\beta}_b$ can be optimized by back-propagation. \hat{Y}_b denotes the regression result. The dice icon refers to randomly sampling a band Y_b as the regression target.

regression approach will allow the model to focus more on the intrinsic association between spectral bands and the overall semantics of the masked image. Experiments show that our S³R-CR and S³R-BR perform better than contrastive learning and masked image modeling methods, and significantly reduces the cost of self-supervised task due to the faster convergence.

2 Method

2.1 Spectral Regression based on Low-rank Prior

Since the spectral correlation in HSIs is extremely high, HSIs can be well represented in a low-dimensional subspace [23]. Let $\mathcal{Y} \in \mathbb{R}^{H \times W \times B}$ denote a HSI presented in 3D tensor with $H \times W$ pixels and B spectral bands. We assume that $\mathbf{Y}_b \in \mathbb{R}^{H \times W}$, denoting the b -th spectral band in \mathcal{Y} , can be represented by a linear combination of the remaining $B-1$ bands with a set of coefficients. This can be formulated as:

$$\mathbf{Y}_b = \boldsymbol{\xi}_b + \sum_{i=1}^B \mathbf{1}[i \neq b] \cdot \mathbf{Y}_i \cdot \beta_{b_i}, \quad (1)$$

where β_{b_i} denotes the linear coefficient corresponding to the i th band, and $\mathbf{1}[\cdot]$ is the indicator function. Let $\boldsymbol{\beta}_b = (\beta_{b_1}, \dots, \beta_{b_{B-1}}, \beta_{b_{B+1}}, \dots, \beta_{b_B})^\top \in \mathbb{R}^{(B-1) \times 1}$. Due to the existence of noise from hyperspectral imaging system, the linear regression process will generate a small error, $\boldsymbol{\xi}_b \in \mathbb{R}^{H \times W}$, which could not be

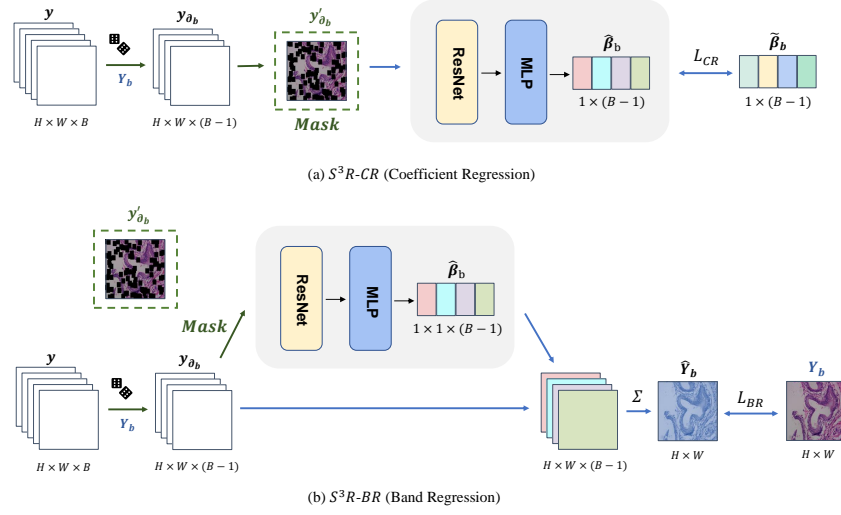


Fig. 2. Our S³R architecture. During pre-training, we randomly select the b th band \mathbf{Y}_b from the input HSI \mathcal{Y} . Then the remaining bands \mathcal{Y}_{∂_b} are masked out by random patches. The encoder is applied to learn a set of coefficients $\hat{\beta}_{b_i}$ from masked bands $\mathcal{Y}'_{\partial_b}$. The learned coefficients are then directly fit $\tilde{\beta}_b$ (a) or reweight \mathcal{Y}_{∂_b} to regress the initially selected band \mathbf{Y}_b (b).

eliminated completely. Eq. 1 provides a linear regression model, which regresses a spectral band by the remaining bands.

To find β_{b_i} , Eq. 1 can be estimated by minimizing the following loss function:

$$\mathcal{L} = \left\| \mathbf{Y}_b - \sum_{i=1}^B \mathbf{1}[i \neq b] \cdot \mathbf{Y}_i \cdot \beta_{b_i} \right\|_F^2. \quad (2)$$

One can simply obtain an estimation of the coefficient $\tilde{\beta}_{b_i}$ through back-propagation by deriving the partial derivatives of β_{b_i} :

$$\tilde{\beta}_{b_i} = \beta_{b_i} - \alpha \frac{\partial \mathcal{L}}{\partial \beta_{b_i}}, \quad (3)$$

or applying a closed-form solution. Let $\tilde{\beta}_b = (\tilde{\beta}_{b_1}, \dots, \tilde{\beta}_{b_{b-1}}, \tilde{\beta}_{b_{b+1}}, \dots, \tilde{\beta}_{b_B})^\top \in \mathbb{R}^{(B-1) \times 1}$. The overall process is shown in Fig. 1.

2.2 Model Based Spectral Regression

Previous section indicates that tensor low-rank prior is an inherent property of HSI which does not rely on any supervisory information. Inspired by the low-rank prior, we propose a Self-supervised Spectral Regression (S³R) architecture for hyperspectral histopathology image classification.

As shown in Fig. 2, we first randomly extract a band \mathbf{Y}_b from an HSI \mathcal{Y} . Let $\mathcal{Y}_{\partial_b} \in \mathbb{R}^{H \times W \times (B-1)}$ denote all bands except the b th band. Next, we randomly mask out \mathcal{Y}_{∂_b} to obtain the masked images $\mathcal{Y}'_{\partial_b}$. Then, a CNN based backbone is used to encode and learn a set of coefficients $\hat{\beta}_{b_i}$ given $\mathcal{Y}'_{\partial_b}$. Last, the learned coefficients $\hat{\beta}_b$ are applied to a pretext task.

In our architecture, we consider two pre-training objectives: (1) coefficient regression (S³R-CR) and (2) spectral band regression (S³R-BR), whose details will be given below.

Image Masking We randomly extract a band $\mathbf{Y}_b \in \mathbb{R}^{H \times W}$ from input HSI $\mathcal{Y} \in \mathbb{R}^{H \times W \times B}$, and mask the remaining bands \mathcal{Y}_{∂_b} with an approximately 65% masking ratio. Like other MIM methods [1], manually adding strong noise interference can significantly make the pretext task more challenging. Note that contents in different bands are masked out independently, which is not the same as MIM.

Coefficients Regression The obtained $\tilde{\beta}_b$ from Eq. 3 can be treated as the groundtruth of the coefficients. Forcing the network to learn the coefficients via randomly masking out a portion of \mathcal{Y}_{∂_b} will make the network to understand the inherent structure of HSI. Residual architecture is proven to work well for HSI classification [14]. Thus, a vanilla ResNet is adopted to encode the masked tensor $\mathcal{Y}'_{\partial_b} \in \mathbb{R}^{H \times W \times (B-1)}$, followed by a MLP head which maps features to the predicted coefficient $\hat{\beta}_b$, where $\hat{\beta}_b = (\hat{\beta}_{b_1}, \dots, \hat{\beta}_{b_{b-1}}, \hat{\beta}_{b_{b+1}}, \dots, \hat{\beta}_{b_B})^\top \in \mathbb{R}^{(B-1) \times 1}$. Noted that the MLP head will not be involved in the downstream task. The loss function for optimizing coefficient regression is computed by:

$$\mathcal{L}_{CR} = \sum_{i=1}^{B-1} \left\| \hat{\beta}_{b_i} - \tilde{\beta}_{b_i} \right\|_1. \quad (4)$$

Spectral Band Regression As mentioned in Sec. 2.1, we assume that one band can be represented as a linear combination of the remaining bands. The selected band \mathbf{Y}_b can be represented by the generated linear coefficients $\hat{\beta}_b$ and \mathcal{Y}_{∂_b} . Then, we aim to minimize the following loss function:

$$\mathcal{L}_{BR} = \left\| \mathbf{Y}_b - \sum_{i=1}^B \mathbf{1}[i \neq b] \cdot \mathbf{Y}_i \cdot \hat{\beta}_{b_i} \right\|_1, \quad (5)$$

where $\hat{\beta}_b$ does not require the supervision of $\tilde{\beta}_b$ learned from Eq. 3.

The two proposed pretext tasks are based on learning the inherent spectral structure, but they have different focuses. S³R-CR tries to figure out the similarity between \mathcal{Y}_{∂_b} and \mathbf{Y}_b , *i.e.*, $\hat{\beta}_b$, and focuses more on the pathological characteristics of different morphologies. S³R-BR could be regarded as a novel MIM method, which focuses more on the holistic semantics of HSIs to recover \mathbf{Y}_b and $\hat{\beta}_b$ is only regarded as a latent variable.

3 Experiments

3.1 Experimental Setup

Datasets We verify the effectiveness of the proposed S³R on two hyperspectral histopathology image datasets, in which all histopathology samples are collected from the gastroenterology department of a hospital. During the process of capturing HSI, the light transmitted from the tissue slice was collected by the microscope with the objective lens of 20 \times . More details are as follows:

PDAC Dataset: It consists of 523 HSI scenes from 36 patients, which are split into 331 for training, 101 for validation, and 91 for testing. Noted that there is no overlap of patients between different splits. Among all the scenes, 255 of them belong to pancreatic ductal adenocarcinoma (PDAC), and the rest ones are normal. The wavelength is from 450 nm to 750 nm, which ends up with 40 spectral bands for each scene. The image size per band is 512×612 .

PLGC Dataset: Clinically, intestinal metaplasia (IM) and dysplasia (DYS) are considered as precancerous lesions of gastric cancer (PLGC) [10] and there are 1105 HSI scenes (414 IM, 362 DYS and 329 normal cases) in the dataset. All samples are randomly split into 661 for training, 221 for validation, and 223 for testing. The wavelength is from 470 nm to 670 nm, which ends up with 32 spectral bands for each scene. The image size per band is 512×512 .

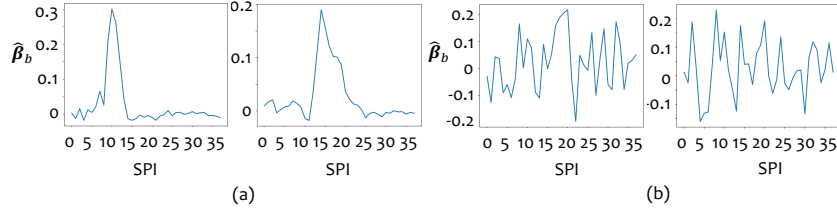
Implementation Details

Pre-Training: The pre-training process of all self-supervised algorithms is conducted on the training set, without any data augmentation. We adopt the ImageNet pre-trained model for all experiments unless otherwise specified. We set the initial learning rate to be 10^{-4} . The maximum training epoch for all the models are 200, and early stop strategy is adopted when training loss no longer drops. We use exponential learning rate decay with $\gamma = 0.99$. The learning rate, maximum training epochs, and learning rate decay for training are the same for other competing methods.

Downstream Task: All pre-trained models can be applied in downstream tasks directly after adjusting the first convolutional layer. During fine tuning, appropriate data augmentation (*e.g.*, scale rotation or gaussian noise) is added to training set. We evaluate the model on validation set after every epoch, and save the parameters when it performs best in current stage. At last, we measure the performance on the test set. Training configurations are consistent throughout the fine-tuning process to ensure fair comparisons. We use AdamW [13] as our optimizer and the learning rate is 10^{-4} with linear decline strategy. The maximum number of fine-tune epoch is 70. All experiments are implemented using Pytorch and conducted on a single NVIDIA GeForce RTX 3090.

Table 1. Ablation study on PLGC dataset.

Image masking	S ³ -BR	S ³ -CR	Backbone	Acc.
×	×	×	ResNet18	94.17
✓	×	×	ResNet18	93.72
✓	✓	×	ResNet18	94.62
✓	×	✓	ResNet18	95.07
×	×	×	ResNet50	95.07
✓	×	×	ResNet50	94.17
✓	✓	×	ResNet50	95.52
✓	×	✓	ResNet50	96.41

**Fig. 3.** Visualizations of $\hat{\beta}_b$. The images in figure (a) come from S³R-CR, and the ones in figure (b) are from S³R-BR. SPI indicates spectral band index. The regression targets from left to right are 10th and 15th band in both figures.

3.2 Ablation Study

The ablations are conducted on PLGC dataset with ResNet and the same pre-trained setup mentioned in sect. 3.1. As shown in Table 1, we first evaluate a straightforward regression method. \mathcal{Y}_{∂_b} is first fed into a vanilla ResNet backbone. Next, a decoder consisting of five transposed convolutional layers is implemented to regress the target band \mathbf{Y}_b . This strategy only obtains 94.17% and 95.07% accuracy with two backbones. While masking the input HSI, the performance gets worse. This may due to the reason that directly regressing the missing band by feeding the remaining bands to a network does not use the inherent structure of HSIs. Thus, it leads to worse performance.

3.3 Comparison between S³R and Other Methods

In this section, we conduct comparisons between S³R and three competitors: 1) contrastive learning (BYOL and SimSiam), 2) self-supervised models designed for remote sensing HSIs (Conv-Deconv [14]) and SSAD [20]), and 3) an MIM-like method [19], termed as Masked HyperSpectral Image Modeling (MHSIM). Vanilla siamese structure based on contrastive learning (*e.g.*, BYOL and SimSiam) is designed for natural images. Thus, we use two strategies to handle the input: randomly selecting 3 bands from HSIs or using all bands. For the MIM-like method, we first randomly mask out a portion of the input, and then reconstruct

Table 2. Performance comparison in Classification Accuracy (%) and Standard Deviation on our two datasets. Scratch means training-from-scratch and IN1K denotes ImageNet-1K. MHSIM is an MIM-like method mentioned in Sect. 3.3.

Method	Dataset	ResNet18	ResNet50
Scratch	PDAC	78.02 (2.28)	80.22 (1.10)
IN1K pre-trained	PDAC	83.52 (1.68)	85.71 (1.68)
BYOL [9] (3 bands)	PDAC	82.42 (1.68)	86.81 (2.28)
BYOL [9] (40 bands)	PDAC	83.52 (1.68)	85.71 (1.68)
SimSiam [4] (3 bands)	PDAC	86.81 (1.10)	86.81 (1.68)
SimSiam [4] (40 bands)	PDAC	85.71 (1.67)	87.91 (1.27)
SSAD [20]	PDAC	80.22 (2.29)	80.22 (2.20)
Conv-Deconv [14]	PDAC	79.12 (1.68)	80.22 (1.68)
MHSIM	PDAC	81.32 (1.10)	83.52 (0.64)
S ³ R-CR	PDAC	91.21 (2.23)	91.21 (1.68)
S ³ R-BR	PDAC	92.31 (1.10)	90.11 (1.27)
Scratch	PLGC	85.20 (1.58)	84.34 (2.87)
IN1K pre-trained	PLGC	93.27 (0.51)	94.62 (1.70)
BYOL [9] (3 bands)	PLGC	94.62 (0.90)	95.07 (0.69)
BYOL [9] (32 bands)	PLGC	94.62 (0.26)	94.17 (0.52)
SimSiam [4] (3 bands)	PLGC	94.62 (0.63)	95.07 (0.90)
SimSiam [4] (32 bands)	PLGC	93.72 (0.69)	92.38 (0.51)
SSAD [20]	PLGC	92.38 (2.99)	91.03 (0.68)
Conv-Deconv [14]	PLGC	90.13 (0.93)	90.13 (0.45)
MHSIM	PLGC	93.27 (1.44)	94.17 (0.52)
S ³ R-CR	PLGC	95.07 (0.21)	96.41 (0.51)
S ³ R-BR	PLGC	94.62 (0.42)	95.52 (0.26)

the missing pixels. Noted that DINO [2] and other transformer-based methods [5] need to be trained on large scale datasets, such as ImageNet. Our dataset contains only hundreds of HSIs. Moreover, due to the lack of computing resources, the batch size for training DINO is 25, which is far from enough. Thus, it is not suitable to compare with transformer-based self-supervised learning methods.

As shown in Table 2, our S³R performs significantly better than MHSIM and other contrastive learning based methods. In particular, on PDAC dataset, S³R-BR outperforms BYOL and SimSiam by 8.79% and 5.5% in classification accuracy with the ResNet18 backbone. On PLGC dataset, S³R-CR with ResNet50 backbone achieves best results. We can observe that, the performance of MHSIM is close to ImageNet pre-training, which is much lower than ours. This may caused by information leakage from CNN architecture in MIM method [7]. Restricted by computing resources, contrastive learning based methods require far more pre-training time (more than 8 minutes per epoch on PLGC) than S³R (about 2 minutes per epoch on PLGC), even with 3-band image as input. Thus, our proposed method can effectively help improve the performance in HSI classification task with lower cost.

We also visualize $\hat{\beta}_b$ in Fig. 3 to further explore our method. We can see that, in Fig. 3 (a), with the coefficient regression, $\hat{\beta}_b$ exhibits a Gaussian-like distribution, which means restoring b th band (peak area in the figure) will be

more dependent on its nearby bands. This may force the model to focus more on the inherent structures of HSIs. As shown in Fig. 3 (b), $\hat{\beta}_b$ learned as latent variable does not present a similar distribution, which illustrates that using pixel-level band regression as the target makes the encoder to extract features by the holistic semantics rather than detailed morphologies.

4 Conclusion

We first attempt to address the problem of self-supervised pre-training for hyperspectral histopathology image classification. We present self-supervised spectral regression (S³R), by exploring the low rankness in the spectral domain of an HSI. We assume one spectral band can be approximately represented by the linear combination of the remaining bands. Our S³R forces the network to understand the inherent structures of HSIs. Intensive experiments are conducted on two hyperspectral histopathology image datasets. Experimental results show that the superiority of the proposed S³R lies in both performance and training efficiency, compared with state-of-the-art self-supervised methods in computer vision.

Acknowledgements This work was supported in part by the National Natural Science Foundation of China under Grant 62101191, and in part by Shanghai Natural Science Foundation under Grant 21ZR1420800.

References

1. Bao, H., Dong, L., Wei, F.: Beit: BERT pre-training of image transformers. In: ICLR (2022)
2. Caron, M., Touvron, H., Misra, I., Jégou, H., Mairal, J., Bojanowski, P., Joulin, A.: Emerging properties in self-supervised vision transformers. In: ICCV (2021)
3. Chen, T., Kornblith, S., Norouzi, M., Hinton, G.E.: A simple framework for contrastive learning of visual representations. In: ICML (2020)
4. Chen, X., He, K.: Exploring simple siamese representation learning. In: CVPR (2021)
5. Chen, X., Xie, S., He, K.: An empirical study of training self-supervised vision transformers. In: ICCV (2021)
6. Devlin, J., Chang, M., Lee, K., Toutanova, K.: BERT: pre-training of deep bidirectional transformers for language understanding. In: Burstein, J., Doran, C., Solorio, T. (eds.) NAACL-HLT (2019)
7. Fang, Y., Dong, L., Bao, H., Wang, X., Wei, F.: Corrupted image modeling for self-supervised visual pre-training. CoRR **abs/2202.03382** (2022)
8. Fu, Y., Zheng, Y., Zhang, L., Huang, H.: Spectral reflectance recovery from a single RGB image. IEEE Trans. Computational Imaging **4**(3), 382–394 (2018)
9. Grill, J., Strub, F., Altché, F., Tallec, C., Richemond, P.H., Buchatskaya, E., Doersch, C., Pires, B.Á., Guo, Z., Azar, M.G., Piot, B., Kavukcuoglu, K., Munos, R., Valko, M.: Bootstrap your own latent - A new approach to self-supervised learning. In: NeurIPS (2020)

10. Gullo, I., Grillo, F., Mastracci, L., Vanoli, A., Fassan, M.: Precancerous lesions of the stomach, gastric cancer and hereditary gastric cancer syndromes. *Pathologica* **112**(3), 166–185 (2020)
11. He, K., Chen, X., Xie, S., Li, Y., Dollár, P., Girshick, R.B.: Masked autoencoders are scalable vision learners. In: CVPR (2022)
12. He, K., Fan, H., Wu, Y., Xie, S., Girshick, R.B.: Momentum contrast for unsupervised visual representation learning. In: CVPR (2020)
13. Loshchilov, I., Hutter, F.: Fixing weight decay regularization in adam. CoRR **abs/1711.05101** (2017)
14. Mou, L., Ghamisi, P., Zhu, X.X.: Unsupervised spectral-spatial feature learning via deep residual conv-deconv network for hyperspectral image classification. *IEEE Trans. Geosci. Remote. Sens.* **56**(1), 391–406 (2018)
15. Salvi, M., Acharya, U.R., Molinari, F., Meiburger, K.M.: The impact of pre- and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis. *Comput. Biol. Medicine* **128**, 104129 (2021)
16. Srinidhi, C.L., Kim, S.W., Chen, F., Martel, A.L.: Self-supervised driven consistency training for annotation efficient histopathology image analysis. *Medical Image Anal.* **75**, 102256 (2022)
17. Wang, J., Li, Q.: Quantitative analysis of liver tumors at different stages using microscopic hyperspectral imaging technology. *Journal of Biomedical Optics* **23**(10), 1 (2018)
18. Wei, C., Fan, H., Xie, S., Wu, C., Yuille, A.L., Feichtenhofer, C.: Masked feature prediction for self-supervised visual pre-training. CoRR **abs/2112.09133** (2021)
19. Xie, Z., Zhang, Z., Cao, Y., Lin, Y., Bao, J., Yao, Z., Dai, Q., Hu, H.: Simmim: A simple framework for masked image modeling. In: CVPR (2022)
20. Yue, J., Fang, L., Rahmani, H., Ghamisi, P.: Self-supervised learning with adaptive distillation for hyperspectral image classification. *IEEE Trans. Geosci. Remote. Sens.* **60**, 1–13 (2022)
21. Zhang, Q., Li, Q., Yu, G., Sun, L., Zhou, M., Chu, J.: A multidimensional choledoch database and benchmarks for cholangiocarcinoma diagnosis. *IEEE Access* **7**, 149414–149421 (2019)
22. Zhao, Z., Wang, H., Yu, X.: Spectral-spatial graph attention network for semisupervised hyperspectral image classification. *IEEE Geosci. Remote. Sens. Lett.* **19**, 1–5 (2022)
23. Zhuang, L., Ng, M.K.: Fasthymix: Fast and parameter-free hyperspectral image mixed noise removal. *IEEE Trans. Neural Networks Learn. Syst.* (2021)